



VERONA UNIVERSITY  
DEPARTMENT OF LINGUISTICS  
SCHOOL OF HUMANITIES  
32<sup>ND</sup> CYCLE

**PhD in Philology, Literature and  
Linguistics**

A COMBINED APPROACH TOWARDS MEASURING  
LINGUISTIC DISTANCE: A STUDY ON SOUTH  
ETHIOSEMITIC LANGUAGES

**Mentor:** Prof. Stefan Rabanus

Signature \_\_\_\_\_

**Coordinator:** Prof. Paolo Pellegrini

Signature \_\_\_\_\_

**Candidate:** Tekabe Legesse Feleke

Signature \_\_\_\_\_

---

# Abstract

The distance among closely related languages is usually measured from three dimensions: structural, functional and perceptual. The structural distance is determined by directly quantifying the phonetic, lexical, morphological and syntactic differences among the languages. The functional distance is measured based on the actual usage of the languages, e.g., intelligibility and inter-lingual comprehensibility. The perceptual distance is related to the subjective judgment of the speakers about the similarity or intelligibility between their native language and the neighboring languages. Studies on language variation measure linguistic distances at least from one of these dimensions. However, as Gooskens (2013), Gooskens & Heuven (2018) and Tang & van Heuven (2009) noticed, languages do not differ just in one dimension; they can be, for example, phonetically similar but syntactically different. The present study, therefore, combined these three perspectives to examine the distance among purposely selected ten South Ethiosemitic languages (Chaha, Endegagn, Ezha, Gumer, Gura, Inor, Kistane, Mesqan, Muher and Silt'e). The study specifically aims to (1) determine the areal classification of the languages; (2) illustrate the similarity or difference between the areal classification of the languages and previous classification by historical linguists; (3) determine the degree of intelligibility among the languages; (4) examine the relationship among the three dimensions of linguistic distances, and (5) explore the major determinants (linguistic and non-linguistic) which contribute to the linguistic distance among the language varieties.

The structural distance was determined by computing the lexical and phonetic differences based on randomly selected 240 words. The lexical distance was defined as the average of pairs of non-cognates in the basic vocabularies. Levenshtein algorithm (Gooskens & Heeringa, 2004; Heeringa, 2004; Kessler, 1995) was used to

compute the phonetic distance. The phonetic distance was defined as an operation that is required to transform a form of sequence of phones. The Semantic Word Categorization test was adapted from Tang & van Heuven (2009) to measure the functional distance. The test involves categorizing words under their semantic categories; e.g., ‘apple’ under ‘fruits’. Self-rating test, based on the recordings of ‘the North Wind and the Sun’, was administered to determine the perceptual distance. With regard to the linguistic determinants, the degree of diffusion of lexical features was estimated using the Neighbor-net network representation and lexicostatistical skewing. The study also examined the influences of some non-linguistic determinants: geographical distance and population size, the degree of contact among the speakers and language attitude. Gabmap was used for clustering and cluster validation. Multidimensional scaling and fuzzy clustering were employed for the cluster validation. The classifications obtained from each of the distance matrices were compared to the previous classifications (by historical linguists) based on the cophnetic distance among various sub-groupings. Two consistency measures (Cronbach’s alpha and Local Incoherence) were employed to derive a combined comprehensive classification of the languages from the three distance measures.

The results of the cluster analysis show that the ten selected South Ethiosemitic language varieties can be fairly grouped into five: {Chaha, Ezha, Gumer, Gura}, {Mesqan, Muher}, {Endegagn, Inor}, {Kistane} and {Silt’e}. This classification is very similar to the classifications previously provided by historical linguists (e.g. Hetzron, 1972). There is also very strong correlation among the measures of the three dimensions of distance. However, these measures have different degree of reliability; the structural distance is the most reliable measure while the perceptual distance is the least reliable distance measure. Furthermore, the results obtained from Word Categorization test show that many of these languages are mutually intelligible. Exceptionally, Silt’e is not mutually intelligible to any of the languages. The results obtained from the analysis of the linguistic determinants show that the similarity among the language varieties is partly the result of the contact among the languages. Moreover, the results of the analysis of the non-linguistic determinants indicate a strong positive correlation between the geographical distance and

linguistics distance, and positive contribution of the contact among the speakers. Nevertheless, there is no significant correlation between the linguistic distance and population size. Besides, among the three dimensions of measuring linguistic distance, it is the perceptual distance that is most affected by the attitude of the speakers.



---

# Acknowledgments

Completing this dissertation would not have been possible without the financial and moral support of several people. Therefore, I would like to express my deepest gratitude to all individuals who have unreservedly helped me in the long journey of completing this dissertation.

First and for most, I am deeply grateful to my esteemed supervisor, Professor Stefan Rabanus. The completion of this dissertation was impossible without his guidance, attentive supervision, encouragement and constant feedback. I am also grateful to my parents, sisters and brothers. I could never cope up the life of solitude in abroad without their encouragement and full material and moral support.

I also gratefully acknowledge Professor Charlotte Gooskens, for the hospitality and kindness with which she welcomed me at the University of Groningen, and for her constant encouragement, feedback and supervision during my stay at the University of Groningen. I would also like to thank Professor Wilbert Heeringa for his invaluable advice and feedback on the application of Gabmap. I am also indebt to Professor Martijn Wieling for his important advice during my stay at the University of Groningen.

My especial thank also goes to Dr. Grirma Mengstu for warmly welcoming me at Addis Ababa University, supervising my stay at Addis Ababa University and creating an opportunity to present my PhD project to the linguistic community at Addis Ababa University. My research stay at Addis Ababa University was enjoyable and fruitful because of his generous collaboration. I would also like to express my heart-felt gratitude to professor Ronny Meyer for the warmly welcome at LLACAN Research Center and for his constant feedback and supervision during my stay at the Center.

I would like to offer my special thanks to Dr. Fekade Menuta for taking time to check the accuracy of the transcription of Gumer words and sending me invaluable comments. I would also like to express my deepest appreciation to Bahire Araya, for sending me alternative transcriptions of some Gura words. My heart-felt gratitude also goes to my colleague Melese Mengesha for spending his priceless time on reading some of the draft chapters of the thesis and sending me very useful comments. I want to also thank my colleagues Mekdes Yikun and Tesfu Teka. Their material and moral assistance during my stay at Addis Ababa University and abroad was especial. My especial thank also goes to the Gurage Zone Tourism Bureau and Silt'e Zone Tourism Bureau. Without the assistance of these two bureaus, my field work could not be easy.

The Universities of Verona and Mekelle University contributed a lot in shaping my academic progress. I, therefore, acknowledge all scholars and the administrative stuffs of those universities who have contributed to my scientific and financial growth. My study abroad was not possible without the financial support by University of Verona. Therefore, I would also want to take this opportunity to thank the University of Verona and the Italian Government. I would also gratefully acknowledge the financial support I received from LLACAN Research Center during my three months' stay at the Center as a guest researcher.

I want to also thank my research assistants in different schools in the Gurage and Silt's zones, especially to Nurahmed Alaw and Kibremariam Nadew (from Arek'it Secondary School); Fikre Shiref and Endalkugn Aklile and Genet (Dinkula Secondary School), Sisay Yirga, Yidnekachew Mulugeta and Solomon Lurga (Agena Secondary School), Gashaw Beharu, Getu Abera and Debebe Teka (Bu'i Secondary School); Yibegeta Gizaw, Amare Sefer, Teka Bedirta (Endebir Seconadry School); Teka Bekana, Jilalu Reshid, Kebede Bereda and Yibegeta Ansa (Yejoka Secondry School); Abdusemed Hassen, Shemsu Nasir and Jilalu (Werabe Secondary School); Melese Tesemma, Semman Asman and Abduzerak Semman (Butajira Secondary School), Temesgen, Mulugeta Degamo and Jilalu Jamal (Gunchire Secondary School); Tigistu Dese, Teseman Gebre, Sileshi Tilahun and Asefa Negash (Hawariyat Secondary School). I know that there are some contributors that I have forgotten

especially from Yejoka Secondary School and Dinkula Secondary school, but your contributions will not be forgotten! Finally, I am regretful to all students in the Gurage and Silt'e zones who spent their time on filling the questionnaires and completing the tests.

---

# Contents

<b>Abstract</b>	<b>i</b>
<b>Acknowledgments</b>	<b>iv</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Stating the Problem . . . . .	1
1.1.1 Research Objectives . . . . .	6
1.1.2 Assumptions . . . . .	6
1.2 Operational Conceptualization of Basic Terms . . . . .	7
1.2.1 Languages, Dialects and Language Varieties . . . . .	7
1.2.2 Linguistic Distance . . . . .	8
1.2.3 Linguistic Similarity . . . . .	8
1.2.4 The Concept of Intelligibility . . . . .	9
1.2.5 Dimensions of Measuring Linguistic Distance . . . . .	9
1.2.5.1 Structural Distance . . . . .	9
1.2.5.2 Functional Distance . . . . .	10
1.2.5.3 Functional Distance vs. Intelligibility . . . . .	10
1.2.5.4 Perceptual Distance . . . . .	11
1.2.6 A Combined Approach . . . . .	11
1.3 Theoretical Framework . . . . .	12
1.4 Expected Outcomes . . . . .	14
1.5 The Organization of the Thesis . . . . .	15
<b>2 Linguistic Features, Geolinguistic and Sociocultural Situations</b>	<b>17</b>
2.1 Ethiopian Linguistic Area . . . . .	17

2.2	The Gurage Area . . . . .	22
2.2.1	The Notion of ‘Gurage’ . . . . .	22
2.2.2	Past Sociolinguistic Situation in the Gurage Area . . . . .	26
2.2.3	Current Demographic and Sociolinguistic Issues . . . . .	28
<b>3</b>	<b>Ethiosemitic Languages</b>	<b>35</b>
3.1	The Origin of Ethiosemitic Languages . . . . .	35
3.2	The Classification of Ethiosemitic Languages . . . . .	39
3.3	Summary: Issues in the Classifications of Ethio-semitic Languages . .	50
<b>4</b>	<b>Measuring Linguistic Distance: the Approaches</b>	<b>53</b>
4.1	Theories of Language Variation . . . . .	53
4.2	Measuring Linguistic Distance . . . . .	56
4.2.1	Structural Distance - Traditional Approach . . . . .	56
4.2.1.1	The Isogloss Method . . . . .	56
4.2.1.2	The Structural Geographic Method . . . . .	57
4.2.2	Structural Distance - Computational Methods . . . . .	58
4.2.2.1	Counting Differences and Similarities . . . . .	58
4.2.2.2	Corpus Frequency Method . . . . .	59
4.2.2.3	Frequency Per Word Method . . . . .	60
4.2.2.4	Levenshtein Distance . . . . .	61
4.2.2.4.1	The Basics of Levenshtein Distance . . . . .	61
4.2.2.4.2	Levenshtein Distance in Previous Studies . .	67
4.2.2.5	Lexicostatistical Approach . . . . .	68
4.2.2.5.1	Basic Concepts of Lexical Distance . . . . .	69
4.2.2.5.2	Lexicostatistical Studies on Non-Ethiosemitic Languages . . . . .	70
4.2.2.5.3	Lexical Distances among Ethiosemitic Lan- guages . . . . .	72
4.2.2.6	Morphosyntactic Computational Approaches . . . . .	78
4.2.3	Functional Distance . . . . .	80
4.2.3.1	Intelligibility and Interlingual Comprehension . . . . .	80

4.2.3.2	Intelligibility in Previous Studies . . . . .	84
4.2.3.3	Intelligibility among Ethiosemitic Languages . . . . .	86
4.2.4	Perceptual Distance . . . . .	91
4.2.4.1	The Notion of Perceptual Distance . . . . .	91
4.2.4.2	Measuring Perceptual Distance . . . . .	92
4.2.4.2.1	Perceived Similarity . . . . .	94
4.2.4.2.2	Perceived Intelligibility . . . . .	95
4.3	Relationship among the Three Dimensions . . . . .	96
4.3.1	The Importance of Examining the Relationship . . . . .	96
4.3.2	Approaches towards Examining the Relationship . . . . .	98
4.4	The Determinants of Linguistic Distance . . . . .	99
4.4.1	Linguistic Determinants . . . . .	99
4.4.1.1	Characteristic Features . . . . .	99
4.4.1.2	Shared Features Due to Contact . . . . .	102
4.4.1.2.1	Estimating Horizontal Diffusion . . . . .	102
4.4.1.2.2	Estimating Borrowed Features . . . . .	103
4.4.1.3	The Influence of Dominant Local Languages . . . . .	107
4.4.2	Non-linguistic Determinants . . . . .	109
4.4.2.1	Contact among the Speakers . . . . .	109
4.4.2.2	Geographical Distance and Population Size . . . . .	110
4.4.2.3	Language Attitude . . . . .	113
<b>5</b>	<b>Research Methods</b>	<b>115</b>
5.1	Assistants and Experimental Subjects . . . . .	115
5.1.1	Research Assistants . . . . .	116
5.1.2	The Experimental Subjects . . . . .	117
5.2	Determining the Structural Distance . . . . .	118
5.2.1	Phonetic Distance . . . . .	120
5.2.2	The Lexical Distance . . . . .	121
5.3	Functional and Perceptual Distances . . . . .	122
5.3.1	Semantic Word Categorization . . . . .	122

5.3.1.1	Materials . . . . .	123
5.3.1.2	The Procedure . . . . .	124
5.3.2	Perceptual Distance and Attitude Tests . . . . .	126
5.3.2.1	The Materials . . . . .	127
5.3.2.2	The Tests and Test Procedures . . . . .	127
5.3.3	Clustering and Cluster Validation . . . . .	129
5.4	Determining Determinants of Linguistic Distance . . . . .	133
5.4.1	Determining Linguistic Determinants . . . . .	134
5.4.2	Determining Extra-linguistic Determinants . . . . .	136
<b>6</b>	<b>Distance among South Ethiosemitic Languages</b>	<b>137</b>
6.1	Structural Distance among the Languages . . . . .	138
6.1.1	Phonetic Distance . . . . .	138
6.1.2	The Lexical Distance . . . . .	141
6.2	Functional Distance among the Languages . . . . .	144
6.3	Perceptual Distance among the Languages . . . . .	148
6.4	The Combined Classification . . . . .	151
6.5	Correlations among the Three Measures . . . . .	157
6.6	Intelligibility among the Gurage Varieties . . . . .	159
6.7	Discussions . . . . .	164
<b>7</b>	<b>Determinants of Linguistic Distances</b>	<b>169</b>
7.1	Linguistic Determinants . . . . .	170
7.1.1	Characteristic Features . . . . .	170
7.1.2	Horizontal Diffusion of Linguistic Features . . . . .	174
7.1.2.1	Network Representation . . . . .	174
7.1.2.2	Estimated Borrowed Features . . . . .	175
7.1.2.2.1	Estimated Borrowed Phonetic Features . . . . .	176
7.1.2.2.2	Estimated Borrowed Lexical Items . . . . .	178
7.1.2.3	The Influence of Local Influential Languages . . . . .	180
7.2	Non-Linguistic Determinants . . . . .	182
7.2.1	Contact among the Speakers . . . . .	182

7.2.2	Geographical Distance and Population Size . . . . .	186
7.2.3	The Attitude of the Speakers . . . . .	189
7.3	Discussions . . . . .	192
<b>8</b>	<b>Theoretical Implications and Future Directions</b>	<b>197</b>
8.1	Intense Contact, Considerable Heterogeneity . . . . .	198
8.2	The Simplicity of the Gravity Model . . . . .	201
8.3	Limitations of the Tree Model . . . . .	204
8.4	The Validity of Non-linguists' Knowledge . . . . .	208
8.5	The Influence of Language Attitude . . . . .	209
8.6	The Combined Approach . . . . .	211
	<b>Bibliography</b>	<b>212</b>
	<b>Appendix</b>	<b>235</b>
<b>A</b>	<b>Data Gathering Tools</b>	<b>236</b>
A.1	Background questionnaire . . . . .	236
A.2	Response sheet for Words Categorization test . . . . .	237
A.3	Response sheet for perceptual and attitude test . . . . .	238
<b>B</b>	<b>Test Materials</b>	<b>239</b>
B.1	List of words for phonetic and lexical distance . . . . .	239
B.2	List of words for Word Categorization test . . . . .	243
B.3	Word Categorization, word order . . . . .	244
<b>C</b>	<b>Supplementary Results</b>	<b>247</b>
C.1	Phonetic similarity index . . . . .	247
C.2	Lexical similarity index . . . . .	247
C.3	Perceptual similarity index . . . . .	248
C.4	Attitude test results . . . . .	248
C.5	Cophenetic distance among the nodes . . . . .	249
C.6	Results of fuzzy clustering . . . . .	250
C.7	Test participants . . . . .	250



---

# List of Figures

2.1	Ethiopian Language Area, from the Language Gulper . . . . .	18
2.2	The Gurage Area, from Meyer (2014) . . . . .	24
2.3	Administrative map of Gurage, from Menuta (2015) . . . . .	29
3.1	Classification of Ethisemitic languages according to Hetzron (1972) .	43
3.2	Classification of Ethisemitic languages according to Demeke (2001) .	48
4.1	Summary of classification of South Ethiosemitic Languages . . . . .	74
4.2	Distribution map: 'traveler' selected as one of the characteristic feature of the Central West Gurage . . . . .	101
4.3	A diagram illustrating the nation of lexicostatistical skewing . . . . .	105
5.1	Procedures of Word Categorization test; S1, S2, S3 etc stands for stimuli, and RT refers to 'Response Time' . . . . .	126
5.2	Example of multidimensional scaling plot with its corresponding map	131
5.3	Example of multidimensional scaling . . . . .	132
5.4	Example of fuzzy clustering . . . . .	133
6.1	Classification of South Ethiosemitic languages based on the phonetic distance . . . . .	140
6.2	Classification of South Ethiosemitic languages based on lexical distance	144
6.3	Classification of South Ethiosemitic languages based on the functional distance . . . . .	147
6.4	Classification of South Ethiosemitic languages based on perceptual distance . . . . .	150

6.5	Comparisons of clusters obtained from the three dimensions of distance	152
6.6	Comparisons between the combined classification and the classifications by historical linguists . . . . .	156
6.7	The mean of the upper and lower half of the intelligibility matrix for the intelligibility among the South Ethiosemitic languages . . . . .	162
6.8	Chaha is selected as a center of communication. The two dimensional arrow shows a symmetry of intelligibility while one-directional arrow shows an asymmetry of intelligibility between pairs of languages. . . .	163
7.1	Language areas based on the pronunciation of words (a) ‘cloud’ and (b) ‘hen’ . . . . .	173
7.2	Neighbor-net dialects, Network-like structure of the data, based on lexical differences . . . . .	175
7.3	Percentage of shared phonetic features between the South Ethiosemitic varieties and the two dominant languages . . . . .	180
7.4	Percentage of shared lexicons between the South Ethiosemitic varieties and the two dominant languages . . . . .	181
7.5	Contact of the speakers of the ten South Ethiosemitic languages. . . .	183
7.6	Degree of contact among the speakers of the ten South Ethiosemitic languages . . . . .	185
C.1	Cophenetic distance . . . . .	249
C.2	Fuzzy clustering based on the structural, functional and perceptual distances . . . . .	250

---

# List of Tables

2.1	Verb paradigms in Amharic . . . . .	20
2.2	Population size in nine districts in Gurage area . . . . .	34
4.1	Phone Frequency Method . . . . .	60
4.2	Examples of Levenshtein distance . . . . .	62
4.3	Examples of Levenshtein distance . . . . .	62
4.4	Feature Representation . . . . .	65
4.5	Percentage of cognates shared among 13 ES languages, Bender (1971)	73
4.6	Percentage of cognates shared among 14 ES languages, Hudson (2013)	75
4.7	Lexical similarity among six Gurage varieties, Menuta (2015) . . . . .	77
4.8	Intelligibility among Gurage varieties, Ahland (2003) . . . . .	88
4.9	Participants' scores on the intelligibility tests by Menuta (2015) . . . . .	90
4.10	Sample lexical similarity matrix . . . . .	106
4.11	Illustration of aggregate skewing of Chaha with respect to the West Gurage varieties . . . . .	106
5.1	Phonetic distance, using Levenshtein algorithm . . . . .	120
5.2	Example of phonetic distance . . . . .	121
5.3	Example of lexical distance . . . . .	122
6.1	Phonetic distance among the South Ethiosemitic languages, the lower half of the matrix, $\alpha = .97$ . . . . .	139
6.2	Lexical distance among South Ethiosemitic languages, lower half of the distance matrix, $\alpha = .87$ . . . . .	142
6.3	Functional distance among the language varieties, $\alpha = .63$ . . . . .	145

6.4	Perceptual distance among the language varieties, $\alpha = .61$ . . . . .	148
6.5	Consistency within the distance matrices . . . . .	153
6.6	Correlation coefficients of the three dimensions of distance . . . . .	158
6.7	Comparison of the correlation coefficients . . . . .	158
6.8	Mean of the participants' score on the Word Categorization test . . .	159
6.9	The adjusted mean of the test-takers' score on the Word Categorization test . . . . .	161
7.1	The pronunciation of ten top ranked words in Central West Gurage varieties (Chaha, Gura, Gumer and Ezha) . . . . .	172
7.2	Ten top ranked words of the West Gurage varieties . . . . .	173
7.3	Aggregate phonetic skewing of the Central West Gurage languages (CWG) with respect to Peripheral West Gurage (PWG) and West Gurage languages(WG). The numbers are percentages. . . . .	176
7.4	Aggregate phonetic skewing of the Peripheral West and West Gurage with respect to other varieties . . . . .	177
7.5	Aggregate lexical skewing of the Central West Gurage languages with respect to Peripheral West Gurage and West Gurage . . . . .	178
7.6	Aggregate lexical skewing of the Peripheral West Gurage and West Gurage language varieties with respect to other varieties . . . . .	179
7.7	The influence of population size and geographical distance . . . . .	187
7.8	Correlation coefficients of phonetic similarity, geographical distance and population size . . . . .	188
7.9	Correlations between speakers' attitude and linguistic distance . . . .	190
7.10	Correlations among the three dimensions of distance and language attitude . . . . .	191

---

# Chapter 1

## Introduction

### 1.1 Stating the Problem

Issues of how to distinguish dialects from languages and how to quantify the resemblance between two or more language varieties have been among the central concerns of dialectology. These two issues are usually addressed by measuring the distance between two or more language varieties. As a general principle, the more two languages are structurally (phonetically, morphologically, lexically or syntactically) similar, the more they are related to each other. If they are similar enough, they are considered dialects of the same language (Tang & van Heuven, 2009, p. 710; Watt, p.2-5). However, distinguishing dialects from languages is more complex than this, and in most cases non-linguistic variables (social, cultural, political, and psychological) play significant roles. This means that determining linguistic distance just based on the structural similarity between the languages may not be sufficient to determine whether two varieties should be considered dialects of a language or two different languages.

In addition to the influences of the non-linguistic variables, there is an inherent limitation of the structure-based approach. The structure-based approach is often criticized for having two drawbacks. First, measuring the linguistic distance requires quantifying the distance among the language varieties. However, languages differ in multiple dimensions (phonology, phonetics, morphology syntax and lexical) and identifying the level that must be measured is a major challenge (Gooskens, 2018,

p. 206; Heeringa et al., 2006, p. 51; Tang & van Heuven, 2007, p. 223; Tang & van Heuven, 2009, p. 710). Second, even if all the levels could be measured, determining the proportion of each level, and squeezing the differences into a single mathematical value is another challenge (Chiswick & Miller, 2005, p .1).

Previous studies of dialectology, in general, have followed two research paths in order to address these limitations. On the one hand, there has been a successful move in terms of shifting from measuring linguistic distance just based on purposefully selected specific linguistic features to measuring distance based on a large aggregate data (e.g., Goebel, 2010; Nerbonne et al., 2011; Nerbonne & Heeringa, 2001; Prokić et al., 2013). On the other hand, different methods that take into account the non-linguistic variables, for example, the perception and the knowledge of native speakers have been developed in the last couple of decades to circumvent the limitations of the structure-based approach (e.g., Preston, 2010). In this regard, the use of mutual intelligibility as a means of measuring linguistic distance and recent advances in folk linguistics have made vital contributions.

As a part of these endeavors, different methods of measuring intelligibility among related languages have been developed (see Casad, 1987; Grimes, 1990; Gooskens, 2013; Gooskens, 2018; Kirk, 1970; Voegelin & Harris, 1951). There have also been various methods of measuring linguistic distance from perceptual perspectives. The perception-based approaches vary in a couple of ways. Some of them examine the perception of the speakers based on carefully selected language inputs such as recorded stories (e.g., Beijering et al., 2008; Gooskens & Heeringa, 2004; Tang & van Heuven, 2009); some others measure the overall perception of the speakers without focusing on a specific language input, for example, by asking in which nearby area a similar language is spoken (e.g., Bucholtz et al., 2007; Montgomery, 2007; Preston, 1996; Pearce, 2009). Likewise, some recent studies focused on examining the perception of non-linguists towards specific sound features such as the features of vowels or consonants (e.g., Labov, 2001; Niedzielski, 1999; Preston, 2010).

Hence, since dialectologists have taken different paths in an attempt to boost the possibility of adequately quantifying the distance among related languages, there has been an immense increase in the methods of measuring linguistic distance. These

methods can be subsumed into three broad categories: structure-based (based on phonetic, lexical or grammatical similarity), function-based (based on inherent and acquired intelligibility) and perception-based (based on the perception of native speakers). Previous studies measured linguistic distance either from one or from the combinations of these three perspectives (Gooskens & Heeringa, 2004, p.196; Tang & van Heuven, 2007, p.223; Tang & van Heuven, 2009, p.710). As noticed by Gooskens (2018, 2018, p.210), the degree of correlation among the linguistic distances measured from each of these perspectives is a concern that requires further exploration.

By examining these three distances, the present study contributes to one of the continuing debates in dialectology, which is to what extent these dimensions of distance correlate. In previous works, there have been doubts, for example, about the reliability of the non-linguists' consciousness in measuring linguistic distance (Goeman, 1999, p.141). The correlation between the degree of intelligibility and degree of linguistic similarity has also been the concern of several recent studies (e.g., Gooskens et al., 2010; Gooskens & Heuven, 2018; Tang & van Heuven, 2007; Tang & van Heuven, 2009). The present study partly indulges these concerns, and examines them in the context of Ethiosemitic languages. In addition to examining the relationship among different perspectives of measuring linguistic distance, the present study aims to determine the distance and degree of intelligibility among selected South Ethiosemitic languages - Chaha, Endegagn, Ezha, Gumer, Gura, Inor, Kistane, Mesqan, Muher and Silt'e.

The above varieties were selected based on two parameters: the number of speakers and the language sub-family they belong to, according to previous classifications by historical linguists. As the study sought to include a high number of participants, language varieties with relatively high number of speakers were selected based on reports in Ethiopian National Statistical Agency (ENSA, 2007). The Ethiopian National Statistical Agency report does not directly provide the number of native Gurage speakers. However, it provides the number of residents in each district. Given that the division of the districts is based on the ethnic background of the residents, the numbers of residents of the ethnically defined districts (e.g., Ezha

District, total population 84,905) was taken as the number of speakers of the varieties. The only exception is Gura, which has a small number of speakers. The inclusion of Gura was based on previous debates among scholars whether it is a separate language or a dialect of Chaha (Hudson, 2013, p.13; Menuta, 2015, p.6). Gura data was obtained based on participants from Gura Megeense and Wirir areas in Chaha district. The study also included at least one language from each of the five groups of the so-called Gurage varieties: Chaha, Ezha, Gumer and Gura (Central West Gurage), Endegagn and Inor (Peripheral West Gurage), Mesqan and Muher (West Gurage), Kistane (North Gurage), and Silt'e (East Gurage).

In addition to measuring the distance among the Gurage varieties, the study aims to examine the major factors that contribute to the distance among the languages. It mainly examines linguistic factors including shibboleths (typical phonetic and lexical features that determine the distance among the languages) and contact-induced features such as borrowing which may have a direct influence on the distance among the languages. The study is also interested in exploring the influences of non-linguistic factors, mainly geographical distance and population size (Trudgill, 1974), contact among the speakers and the attitude of the speakers. Previous studies consistently reported the influence of geographical distance on the linguistic distance. However, the nature of the relationship has remained the issue of academic debate. Trudgill (1974) predicted a positive relationship between the linguistic distance and the square of the geographical distance. In other words, the relationship is quadratic, according to Trudgill's Gravity model. Several later studies reported a sub-linear relationship between the two variables (e.g., Gooskens & Heeringa, 2004; Nerbonne et al., 2005). The relationship between population size and linguistic similarity is similarly controversial. Some studies reported some degree of relationship between the two variables (e.g., Heeringa et al., 2011) while others reported the complete absence of the relationship (e.g., Nerbonne et al., 2005).

Several previous studies recognize the influence of contact among the speakers on the degree of linguistic similarity (e.g., Axelsen & Manrubia, 2014; Gooskens & Van Bezooijen, 2006; Gooskens & Heuven, 2018). Nonetheless, often mixed results are reported regarding the influence of the speakers' attitude on the distance



among the languages, some studies reported the presence of the influence of language attitude (Abu-Rabia, 1996; Golubović & Sokolić, 2013; Pavlenko, 2006); others reported the marginal role of language attitude in determining the linguistic distance (Schüppert & Gooskens, 2011; Van Bezooijen & Gooskens, 2007). The present study examined the influences of these factors in the context of South Ethiosemitic languages.

Ethiosemitic languages are Semitic varieties which are spoken in Ethiopia and Eritrea. They are broadly divided into North and South Ethiosemitic. The South branch is the focus of the present study. South Ethiosemitic varieties are spoken in central, east and southwestern part of Ethiopia. The present study specifically focuses on language varieties which are traditionally called ‘Gurage languages’- they are spoken in a tiny area in the southwest of Ethiopia. I frequently used the term ‘South Ethiosemitic Languages’ rather than ‘Gurage Languages’ because of two reasons. First, there is no agreement among Semitists whether the so called Gurage languages refer to a single genetically attested unit (Hetzron, 1972, p. 119; Meyer, 2011, p.1221). Second, some of the speakers of these varieties do not consider themselves as Gurage - Silt’e is a good example. According to Meyer (2011, p. 1223), only Sebat Bet Gurage ‘Seven House Gurage’ identify themselves as Gurage. The traditional ‘Gurage languages’ refer to several varieties among which some are unstudied. In spite of several attempts (e.g. Demeke, 2001; Hetzron, 1972; Hetzron, 1977; Leslau, 1967; Leslau, 1969) to classify these language varieties, there is yet a great number of language varieties whose exact areal and genealogical position is still debatable (e.g. Mesqan, Muher and Inor). A part of the problems is that the classifications of the languages are often based on a small number of grammatical, morphological and phonological shared features since detailed descriptions of some of the languages are not available (Demeke, 2001, p.67; Hetzron, 1972, p.1).

The lack of detailed evidence, combined with other factors such as a long history of contact among Ethiosemitic and other neighboring Afroasiatic languages, compelled previous studies to draw often sketchy conclusion regarding the origin and the classification of the languages (Goldenberg, 1977, p.462). So far, there is no clear well-agreed proposal about the origin and the classification of Ethiosemitic lan-

guages in general (Demeke, 2001, p.59; Goldenberg, 1977, p.462; Hetzron & Bender, 1976, p.3; Hudson, 2000, p.79; Hudson, 2013, p.20). Furthermore, although there are some studies for example by Gutt (1980), Ahland (2003) and Menuta (2015), on the intelligibility of these language varieties, the degree of intelligibility among many of the language varieties has not been thoroughly investigated. The overwhelming claims about the intelligibility of some of the language varieties are intuitive suggestions by scholars who are rather interested in the structural differences among the languages (Demeke, 2001, p.81; Hudson, 2013, p.18).

### 1.1.1 Research Objectives

The present study aims to address two general objectives: (a) methodological and (b) linguistic. There are two specific objectives related to the methodological part; (1) determining to what extent the three dimensions of linguistic distance correlate; (2) examining to what extent the taxonomies obtained from structural, functional and perceptual distance measures are similar to the classifications previously provided by historical linguists. There are four specific objectives related to the linguistic part, i.e., the South Ethiosemitic language varieties; (1) determining the distance among the selected language varieties; (2) determining the degree of intelligibility among the varieties; (3) classifying the languages using the data obtained from the three dimensions of distance; and (4) identifying factors that are contributing to the distance among the selected language varieties.

### 1.1.2 Assumptions

With regard to the distance among the South Ethiosemitic languages, the present study presupposed a close similarity among many of the language varieties, and that this similarity could be due to the influence of non-linguistic factors such as geographical distance, population size and contact among the speakers. The assumption about the similarity among the languages was based on previous studies which were devoted to the classification of South Ethiosemitic languages such as Demeke (2001), Hetzron (1972), Hetzron (1977), and Hudson (2018). The assump-

tion about the influence of non-linguistic factors was based on previous studies on non-Semitic languages such as Nerbonne (2010), Nerbonne et al. (2011) and Trudgill (1983). The study also expected a high degree of intelligibility among some of these language varieties based on previous reports by Ahland (2003), Gutt (1980) and Menuta (2015). Regarding the methodological concerns, the study assumes strong correlations among the three dimensions of distance, and very close similarity between the anticipated areal classification and the classifications previously proposed by historical linguists. These predictions were based on the results previously reported in the studies which were conducted on non-Semitic languages such as Chiswick & Miller (2005), Gooskens (2018), Nerbonne & Heeringa (2001) and Tang & van Heuven (2009), and studies on the similarity among Ethiosemitic languages (e.g., Bender, 1971; Bender et al., 1972; Leslau, 1969; Menuta, 2015).

## 1.2 Operational Conceptualization of Basic Terms

This section presents the definitions of key terms that are used throughout the present study. Some of these terms may have different interpretations in some other studies. Unless they are presented with the citation of other sources, the definitions provided in this section are binding for the interpretations of ideas presented in this study.

### 1.2.1 Languages, Dialects and Language Varieties

The difference between language, dialect and language variety can sometimes be confusing. In some studies, both language and dialect refer to a regional variety (Charles et al., 2018, p.4-5; Tang & van Heuven, 2019, p.709), and the distinction between the two is usually based on non-linguistic factors such as politics and speakers' attitude (Tang & van Heuven, 2009, p.709). For instance, some Chinese languages are considered 'dialects' though they are not mutually intelligible; in contrast, there are many mutually intelligible Indian languages that are considered separate languages. In most cases, however, two or more language varieties are considered dialects if they are consistently proved to be mutually intelligible; otherwise they are considered as

distinct languages (Britain, 2018, p.143; Gooskens, 2018, p.205; Tang & van Heuven, 2009, p.709; Watt, 2018, p.5). This conception has also its own shortcoming since there is a problem of asymmetry of intelligibility. Regardless of these limitations, the present study adopted this later conceptualization; if two or more language varieties are consistently proved to be intelligible, they are dialects of the same languages. In the present study, the term ‘language variety(ies)’ is used in two conditions; in a situation where there is no clear evidence of intelligibility or unintelligibility, and in contexts where both intelligible and unintelligible languages should be collectively referred.

### **1.2.2 Linguistic Distance**

Linguistic distance refers to how different one language variety is from another (Chiswick & Miller, 2005, p.1). Linguists usually measure linguistic distance based on structural parameters such as lexical, phonetic and morpho-syntactic difference among related languages. In the present study, linguistic distance represents a broader sense; in addition to the structural differences, it also includes differences between related languages because of the difficulty of understanding of the speakers (functional distance) and perception of the speakers (perceptual distance).

### **1.2.3 Linguistic Similarity**

Linguistic similarity refers to how similar one language variety is to another. It is just the reflection or the mirror image of the linguistic distance. In terms of interpretation, there is no significant difference with regard to the linguistic distance between two related languages and the linguistic similarity between the same two related languages. For instance, 20% lexical similarity between any given pair of languages, let say language ‘A’ and language ‘B’, is exactly the same with 80% lexical difference between the two hypothetical languages.

### 1.2.4 The Concept of Intelligibility

Intelligibility is defined as the degree to which speech variety ‘A’ is understood by speakers of speech variety ‘B’ mainly because of the historical relationship between speech varieties ‘A’ and ‘B’ (Gutt, 1980, p.58). Gutt (1980, p.57) makes a distinction between intelligibility and inter-lingual comprehension; inter-lingual comprehension is the degree to which a speech variety ‘A’ is understood by speakers of a speech variety ‘B’ or vice versa due to contact between the two languages. Therefore, inter-lingual comprehension specifically refers to the understanding between speakers of two or more varieties not only because of the historical similarity between the languages but also because of inter-lingual learning. Similarly, Gooskens & Heuven (2018, p.2-3) makes a distinction between inherent intelligibility and acquired intelligibility. As the word implies, ‘inherent intelligibility’ is the degree of communication between the speakers of certain languages mainly due to the historical relatedness between the languages while ‘acquired intelligibility’ is strongly related to language learning. In the present study, mutual intelligibility refers to the inherent intelligibility. ‘Mutual’ in mutual intelligibility implies bi-directional understanding (symmetric). Since this is not always the case, sometimes scholars distinguish between bi-directional intelligibility (symmetric) and uni-directional intelligibility (asymmetric) (e.g., Gooskens et al., 2010, p.4; Gooskens, 2018, p.213; Tang & van Heuven, 2009, p.710).

### 1.2.5 Dimensions of Measuring Linguistic Distance

In the present study, *dimension* refers to perspectives from which the linguistic distance is measured. Linguistic distance is generally measured from three perspectives: structural, functional and perceptual. The distance values obtained from these measures are structural, functional and perceptual distances respectively.

#### 1.2.5.1 Structural Distance

The structural distance is often considered as a real linguistic difference. It is measured by determining the phonetic, phonological, morphological, syntactic or lexical

differences among related language varieties (Tang & van Heuven, 2009, p.713). Many of the previous studies in comparative linguistics, dialectology and dialectometry have measured linguistic distance based on the structural differences among related languages.

### **1.2.5.2 Functional Distance**

In the present study, functional distance (Tang & van Heuven, 2009, p.712) refers to the degree of difficulty of understanding between the speakers of two or more related languages. In other words, functional distance is a reflection or a mirror image of the degree of mutual intelligibility. It is called ‘functional distance’ since it is measured based on the actual communication or understanding between the speakers of related languages which is the fundamental function of human language.

### **1.2.5.3 Functional Distance vs. Intelligibility**

In the present study, for the sake of expediency, functional distance and degree of intelligibility are used to express very related but slightly different concepts. As indicated above, intelligibility refers to the degree of ease of understanding or communication whereas functional distance refers the degree of difficulty of understanding among the speakers of the languages. This distinction is important for logical reasons. First, as indicated above, often a distinction is made between inherent intelligibility and acquired intelligibility. Even, for some, only inherent intelligibility is considered as an intelligibility (e.g., Gutt, 1980; Tang & van Heuven, 2009). Therefore, ‘functional distance’ can be used as a general term which refers to the linguistic distance that is measured based on either the inherent intelligibility or acquired intelligibility or both. Second, both inherent intelligibility and acquired intelligibility are parts of the actual communication - which is the main function of the language. Hence, functional distance (function-based distance) can best describe a distance which is measured from this perspective. More importantly, using functional distance, helps to show a distinction between the intelligibility which is measured based on the actual performance and perceived intelligibility, which is measured based on the perception of non-linguists. The functional distance refers

only to the former one.

#### 1.2.5.4 Perceptual Distance

Perceptual distance is often called *speakers' consciousness or perceived distance* (Kremer, 1999, p.32; Preston, 1999, p.xxiii; Sibata, 1999, p.40; Long, 1999, p.199; Mase, 1999, p.101). It is what the speakers think that they understand, without giving proof of their degree of understanding. It can be a perception about certain local language varieties in general or about a specific speech sample such as recorded speech. Perceptual distance can be either perceived similarity or perceived intelligibility (Tang & van Heuven, 2009, p. 712-713). Perceived similarity is what the speakers think about the similarity between their native languages and other related local languages. Perceived intelligibility is related to the perception of the speakers about the degree of intelligibility between their native language and other related local languages. In the present study, 'perceptual distance' is the combination of perceived similarity and perceived intelligibility.

#### 1.2.6 A Combined Approach

It was indicted in 1.1 that measuring every level of the structural (phonetic, morphological, syntactic...etc) differences or similarity is impossible. Hence, functional distance is usually considered as an alternative (Tang & van Heuven, 2009, p.273) to the structural distance since intelligibility tests inherently contain all language features: phonetic, morphological, syntactic... etc. Nevertheless, constructing and administering intelligibility tests are usually difficult, especially if several language varieties must be tested. Determining the intelligibility threshold is also a challenge (Gooskens, 2018, p.206; Menueta, 2015, p.38). Moreover, asymmetric nature of intelligibility makes it somehow an unreliable tool. Speakers' consciousness about their native language and about other surrounding varieties is often taken as another perspective of measuring linguistic distance (Gooskens, 2018, p.212-215; Schüppert & Gooskens, 2011, p.122; Preston, 1999, p.177-203). However, the affinity of perceptual distance with the speakers' attitude and with other non-linguistic variables makes it the least reliable approach (Goeman, 1999, p.141). Hence, in order to cir-

cumvent these limitations, the combination of structural, functional and perceptual distance were applied to the South Ethiosemitic languages. This combination is called ‘a combined approach’.

### 1.3 Theoretical Framework

The present study relies on the notions of Wave model and recent models of language contact such as Gravity model (Labov, 2001; Trudgill, 1974), Dynamic model (Schneider, 2003), and Dialect Formation model (Trudgill, 2004). These models presuppose that language change is the outcome of the interaction between linguistic features and non-linguistic variables. The Wave model assumes that language features spread as concentric wave affecting the immediate area before the remote area. The spread of features is related to the geographic distance of the varieties; the closer the features are to each other, the higher their degree of affectedness. In the Gravity model (Labov, 2001; Trudgill, 1974), the diffusion is hierarchical in a sense that the features diffuse from cities (center) to the low populated rural areas. It further assumes that the diffusion of linguistic features is not always continuous; remote big centers can be affected before small intermediate areas. Hence, the diffusion of linguistic features is determined not only by territorial proximity, but also by various other non-linguistic factors. The Dynamic model assumes that dialect formation is a gradual multi-step process. It is determined by non-linguistic factors such as geographical proximity, ethnicity, social integration such as intermarriage and other demographic factors. In other words, geographical distance, population size, language attitude and other social variables play a significant role in the process of new dialect formation (also see Britain, 2018; Nerbonne & Heeringa, 2001; Spruit, 2006; Tang & van Heuven, 2009). The present study takes some of these variables (contact among the speakers, geographical distance, population size and the attitude of the speakers) into account. These variables have been identified as major non-linguistic factors that affect linguistic distance (see Heeringa et al., 2011; Gooskens, 2005; Gooskens, 2013; Menuta, 2015; Nerbonne et al., 2005)

To effectively employ the proposed combined approach, the structural distances



among the language varieties were computed using lexicostatistics and Levenshtein algorithm based on a randomly selected 240 list of words. The perceptual distance (perception-based distance) was determined using subjective test that requires the judgment of the speakers about their understanding of recorded speeches of their own native languages and other neighboring language varieties. For the functional distance measure, Semantic Word Categorization test was adopted from Tang & van Heuven (2009). In this test, the selected participants were requested to categorize a list of words under various semantic categories; for example, ‘apple’ was categorized under ‘fruits’. Gabmap was used for the computation of Levenshtein distance, for the clustering and cluster validation. Multidimensional scaling and fuzzy clustering were employed to validate the clusters. Gabmap was also employed to determine the specific features (shibboleths) of the main language areas, and to create the dialect map of the language areas. The hypotheses of the Gravity model (Trudgill, 1974) were examined using travel time (time that a public transport takes to travel from one language site to another) and travel distance (the estimated distance between each study site) between language sites, and the population size of the speakers of the languages of interest was obtained from ENSA (2007).

The present study relies on methods of dialectometry (Goebel, 1982; Goebel, 2010; Heeringa, 2004; Nerbonne & Kretzschmar, 2003; Séguy, 1973) to determine the structural distance among the South Ethiosemitic languages. Dialectometry is a quantitative and computational branch of dialectology. It focuses on the general relationship among languages, not on specific details which is often the case in dialectology. In other words, dialectometry aggregates linguistic data to quantify the distance among related languages. Hence, the fundamental motivation of dialectometry lies in the opportunity to condense a large amount of data. The possibility to aggregate a large amount of data further provides an opportunity to deploy various statistical analysis.

## 1.4 Expected Outcomes

The present study has methodological, linguistic and pedagogical advantages. From methodological point of view, the comparison of the three dimensions of distance provides insight about the reliability of the perception and function-based linguistic distance measurements. Moreover, by examining the relationships among the three dimensions of linguistic distance, the study suggests a combined approach which minimizes the weaknesses of the structural perspective. If a strong correlation is found among the three dimensions, it would also be possible to avoid the time and energy that could be required for the areal classification of the languages based on grammatical/syntactic descriptions. In other words, for example, using the degree of mutual intelligibility as a criterion for classification could be considered to measure the linguistic distance in the situations where linguistic data is not available. If it is consistently proved that the perception of the speakers is a reliable linguistic data, it could be considered as an option to get ride of the difficulty of constructing intelligibility tests and the constraints related to administering the tests. The correlation results may pave a way for a significant methodological shift in the study of language variation; strong correlation implies the possibility of substitution between functional, structural and perceptual distances. This means there will be the possibility for future studies to unreservedly employ anyone of the three perspectives to determine the similarity and intelligibility among closely related languages.

From a linguistic point of view, the present study positively contributes to the enduring controversy about the classification of Ethiosemitic languages in general, and the South Ethiosemitic languages in particular. Though the present study never tends to make a claim about the genealogical relationship among the languages, it surely provides evidence that is useful to understand the dynamics within the Ethiosemitic languages. Moreover, the present study provides linguistic evidence that can be utilized for pedagogical purposes. All the languages under discussion are minority languages. Except Silt'e, all are spoken languages; there are proposed writing systems (for example for Chaha and Kistane) but they have not been used in schools and other settings. One of the main challenges of using the Gurage varieties

for schooling and administration is the lack of successful effort to standardize the languages (Meyer, 2018). Hence, identifying the languages that are similar and those that are different provides valuable inputs for the ongoing standardization efforts.

## 1.5 The Organization of the Thesis

The present study is classified into eight chapters. Following this introduction, Chapter two describes the geolinguistics and sociocultural settings in which the language varieties of the present study situated. This chapter discusses four major points: the general overview of Ethiopian linguistic area, the notion of ‘Gurage’, past sociolinguistic situation in the Gurage area, and demographic and social issues of the Gurage people. Chapter three is about Ethiosemitic languages. It presents the hypotheses about the origin of Ethiosemitic languages, the classification of Ethiosemitic languages and enduring challenges in the classification of Ethiosemitic languages. Chapter four deals with the theoretical and methodological issues related to measuring linguistic distance. The chapter presents the details of the structural, functional and perceptual distance measures, and methods that have been used to measure these distances. The discussion of the methods of measuring structural distance is restricted to the computational approach.

Chapter five presents the experimental part of the present study, focusing mainly on the sampling procedures, descriptions of the types of tests employed, the tasks designed to measure the distance and the degree of intelligibility among the languages, and data collection and administration procedures. Chapter six presents results of the measurements among the South Ethiosemitic languages. The results of the structural, functional, and perceptual measurements among the selected languages are presented based on the data obtained using the methods presented in Chapter five. This chapter also presents the results of the relationship among the three dimensions of linguistic distance. Chapter seven presents results of the analysis on the determinants of the linguistic distance. In this chapter, linguistic and non-linguistic factors that determine the distance among the South Ethiosemitic languages are presented. Finally, Chapter eight presents the theoretical implication

of the distance among the South Ethiosemitic languages and the factors contributing to the distance among the languages.

The contents presented in these chapters can be grouped into three broad conceptual categories. Chapter one, two, three and four cover the conceptual part. Chapter five, six and seven deal with the empirical part of the study. Chapter eight presents the theoretical implications of the empirical facts presented in Chapter five, six and seven.

---

## Chapter 2

# Linguistic Features, Geolinguistic and Sociocultural Situations

This chapter presents the general overview of Ethiopian linguistic area (§2.1), the territories in which the ten language varieties are spoken (§2.2.1), past sociolinguistic situation in the Gurage area (§2.2.2), and current sociolinguistic and socio-economic issues associated to the speakers of the language varieties under investigation (§2.2.3). By elaborating on these issues, the present chapter provides the background that is essential to understand the current and the past linguistic and social metamorphosis in the Gurage land.

### 2.1 Ethiopian Linguistic Area

The Gurage area is encapsulated within the Ethiopian linguistic area or *Sprachbund*, one of the most linguistically diverse places (Crass & Meyer, 2009, p.1; Ferguson, 1976, p.63 Tosco, 2000, p.329; Zaborski, 1991, p.123). The Ethiopian linguistic area encompasses a large part of east Africa including Ethiopia, Eritrea and Djibouti. The scope of the area may stretch even more beyond this territory (Crass & Meyer, 2009, p.2). This area can be considered as ‘a melting pot’ where languages of four different language families: Omotic, Nilo-Saharan, Cushitic and Semitic are intermingled through a long history of contact. Though the area has received the interest of linguists for decades, the diffusion of features within languages of this area has

not been thoroughly investigated. This is partly because the detailed descriptions of the features of some of the languages are not available. Relatively, the shared features of the Semitic and major Cushitic languages such as Oromo have received some degree of attention, e.g., Appleyard (2015), Leslau (1945) and Leslau (1952). However, there is yet a lot to be known about the contact between several Cushitic and Ethiosemitic languages. Indeed, the intermingling between Semitic and Cushitic features is one of the challenges that linguists have faced in the attempt to classify the Ethiosemitic languages (Appleyard, 2015, p.16; Hudson, 2000, p.79; Goldenberg, 1977, p.499). Different features of languages in the Ethiopian linguistic area have been identified based on surveys on some selected languages e.g., Crass & Meyer (2009) and Ferguson (1976). For the sake of illustration, features presented in two studies, i.e., Ferguson (1976) and Crass & Meyer (2009) are summarized below. Ferguson (1976) is still considered as a main reference by many. Crass & Meyer (2009) is relatively a recent survey on the features of the languages of this linguistic area.

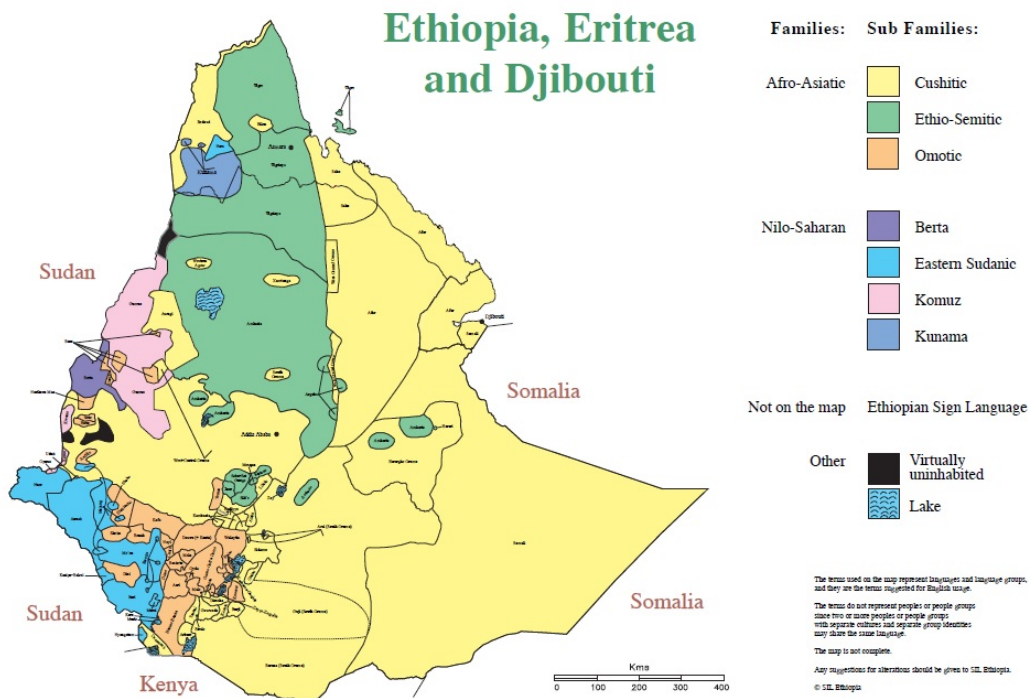


Figure 2.1: Ethiopian Language Area, from the Language Gulper

Ferguson (1976) identified eight phonetic and 18 morphological and grammatical features of eighteen selected languages, which characterize the Ethiopian linguistic area. He argued that languages of Ethiopia display features that are peculiar to the area. Some of these shared features are due to genetic inheritance while others are the outcomes of reciprocal diffusion. The identified phonological features include the replacement of /p/ by /f/ as a counterpart of /b/, the palatalization of dental consonants as a common grammatical process, /g, s, c, n/ as /ǰ, š, č, ñ/, varieties of ejectives including /p', t', c', k', s', p'/, the occurrence of an implosive /d/, the pharyngeal fricatives /ħ, ʕ/, contrastive consonant germination as in Amharic *wana* 'swimming' vs. *wanna* 'principal', relatively short central vowels /i/ and /ə/, and the presence of an epenthetic vowel /i/ as in Amharic *isport* 'sport', *dīrama* 'drama'.

The eighteen morphological and syntactic features are (1) SOV word order; (2) postpositions as for example Amharic *lay* in *igir lay* 'on a leg', *wist* in *bet wist* 'in a house' and *ǰərba* in *kəne ǰərba* 'behind me'; (3) possessive suffixes are very similar to object suffixes added to the verb as in Amharic, *bet-accəw* 'their house' vs *gəddəl-ə-accəw* 'he killed **them**'; (4) Root and pattern morphology i.e., the basic meaning of the word is encoded in the roots which are constituted by consonants, and the grammatical features are expressed by the vocalic elements - vowels inserted within the consonants. Based on the pattern of the consonants and the vocalic elements, verbs are usually classified into three: type 'A', 'B', and 'C'. Type 'A' verbs have geminated consonant clusters at penultimate position only in the perfective aspect. In type 'B' verbs, the penultimate consonants are geminated in all word forms while type 'C' verbs contain a vowel /a/ after the first consonant. In the Amharic examples in Table 2.1, *yi-* is a subject prefix, and *-ə* is person, gender and number suffix.

Table 2.1: Verb paradigms in Amharic

Paradigms	Type A	Type B	Type C
Perfective	<i>səbbər-ə</i> ‘broke’	<i>fəlləg-ə</i> ‘found’	<i>k’arrəm-ə</i> ‘collected’
Imperfective	<i>yi-səbr</i> ‘breaks’	<i>yi-fəllig</i> ‘finds’	<i>yi-k’arrim</i> ‘collects’
jussive	<i>yi-sbər</i> ‘shall break’	<i>yi-fəllig</i> ‘shall find’	<i>yi-k’arm</i> ‘shall collect’
Imperative	<i>sibər</i> ‘break!’	<i>fəllig</i> ‘find!’	<i>k’arrim</i> ‘collect!’
Infinitive	<i>mə-sbər</i> ‘to break’	<i>mə-fəlləg</i> ‘to find’	<i>mə-k’arrəm</i> ‘to collect’
Gerundive	<i>səbr-o</i> ‘breaking’	<i>fəllig-o</i> ‘finding’	<i>k’arm-o</i> ‘collecting’

Another feature of this area is that (5) in complex sentences, subordinate clauses usually precede the main clauses. (6) The use of converb constructions is another feature. Converb constructions have various grammatical purposes in these languages (see Banti, 2010; Hetzron, 1972, p.98-115; Wetter, 2007; Völlmin et al., 2007). (7) Quoting clauses - according to Ferguson (1976), languages of this area have “a frequent construction clause which purports to be a direct quotation followed by a form of the verb to say”. (8) Compound verbs are also common in these languages; i.e., the combinations of verbs or adjectives with the verb *alə* ‘to say’ form compound words that have nothing to do with ‘saying something’ as in Amharic examples: *k’u’éc alə* ‘to sit down’ *bidigg alə* ‘to stand up’ *k’əs alə* ‘to slow down’ (see Amha, 2010; Amberber et al., 2010). (9) Negative copula - negative copulas have either different stem or different negation marking suffixes, compared to non-negated counterpart. In Amharic, for example, the positive copula *nəw* ‘is’ negated as *ay-dəllə-m* ‘it is not’, and *allə* ‘exist’ is negated as *yəlləm* ‘does not exist’.

(10) In noun phrases of some of the languages of this area, numerals assign singular number to nouns, not plural numbers as in Amharic: *amist səw* ‘five man’, *məto fyyəl* ‘hundred goat’, *ših bet* ‘thousand house’. (11) Reduplicated intensives as in Amharic *k’orrət’ə* ‘cut’ vs. *k’orarrət’ə* ‘cut into tiny pieces’, *səbbərə* ‘break’ vs. *səbabbərə* ‘break into pieces’. (12) Broken plurals is another peculiar feature of languages of this area, i.e., number is expressed by the change of patterns<sup>1</sup> of consonants and vowels within the singular form rather than by the concatenation of

<sup>1</sup>In these two particular examples, there is a change of pattern from CVCCVC to CVCVCVC





Singular form must be marked, *re?e-tti* ‘a female goat’ or ‘a small goat’. The final suffix marks both feminine gender and diminutive. The masculine singular is *re?e-čča* ‘male goat’.

Crass & Meyer (2009) is a recent study on the features of Ethiopian linguistic area. This study provided morphological and syntactic features, many of them which are not included in Ferguson (1976). Eight languages were surveyed in this study: Gumer, Muher, Wolane, Zay, Amharic, Libido, Oromo and Kabeena. The first five are Ethiosemitic languages while the last three are Cushitic languages. There are many other studies on Ethiopian linguistic area; for example, Bender (2003) on Nilo-Saharan languages using Ferguson’s parameters, and Zaborski (1991) on selected Semitic and Cushitic languages. Not all scholars agree on the concept of an Ethiopian linguistic area though. The disagreements seem to have two categories. Some recognize the presence of an Ethiopian linguistic area, but disagree on the type of features used and the number of languages that should have been compared (e.g., Zaborski, 1991), some others question the existence of the Ethiopian linguistic area claiming that many of the features are genealogically inherited (eg., Güldemann & Fiedler, 2019; Tosco, 2000). In general, the studies seem to suggest two essential points; the complexity of the Ethiopian linguistic area and the lack of consensus among scholars on the basic definition of ‘linguistic area’. Inconsistency in the conceptualization of ‘linguistic area’ is extensively discussed in Bisang (2006).

## 2.2 The Gurage Area

This section presents three major issues that are related to the Gurage area; the notion of Gurage (§2.2.1), past sociolinguistic situations in the Gurage area (§2.2.2) and current demographic, socioeconomic and sociolinguistic issues (§2.2.3).

### 2.2.1 The Notion of ‘Gurage’

The language varieties examined in the present study are spoken in the southwest of Ethiopia, about 160 km from Addis Ababa. This area is traditionally called *Gurage area*. According to Menuta (2015, p.4) and Meyer (2011, p.1221), the Gurage area

stretches from south of Welkit'e in the west (8°17.68 N and 37°47.20 E) and Bu'i in the north (8°19.59 N and 38°33.03 E) to a few miles south of Dinqulla in the east (7°52.15 N and 37°48.50 E) and Qabul in the west (7°52.58 N and 38°02.02 E).

The term *Gurage* is also traditionally used to express both the linguistic area and the people living in the area (Menuta, p.15; Meyer, 2011, p.1221). However, some of the communities in this territory do not identify themselves as Gurage. As indicated in 1.1, Silt'e people are among them. Indeed, some of the communities in the Gurage area are non-Semitic speakers, for example, Libido and K'abeena. Strictly speaking, the term *Gurage* refers neither to a single language nor to a common historical or cultural unit (Meyer, 2011, p.1221). The Gurage area is linguistically extremely diverse; in most cases different language varieties are further divided into dialects and sub-dialects. In some cases, names of the languages and of the clans are intermingled (Hetzron, 1972, p.3; Hetzron, 1977, p.4).

A mystery behind this diversity in this compact territory where there is no significant geographical barriers that prevent contact between neighboring tribes has remained undisclosed. Hetzron (1977, p.4) indicated that 'clannish separatism' in the past is the main cause of the surface differentiation among the Gurage varieties. It seems that previous history of war, tribe-oriented social division and ethnic-based political systems played their own roles. There were also external forces. It was indicated in 2.1 that all Ethiosemitic languages, in general, have a long history of contact with other Afroasiatic languages that are spoken in Ethiopian (see Figure 2.1). It seems that the impact of the contact between the Semitic and Cushitic languages increases as one moves from the north to the south. This is due to the fact that the southern part of Ethiopia was inhabited by a large number of non-Semitic speakers, predominantly Cushitic and Omotic speakers. Hence, being the south pocket of the Ethiopian Semitic, the Gurage varieties are the most influenced ones, compared to other Ethiosemitic languages.

According to Menuta (2015, p.5), Tesema et al. (1993, p.5) and Tilahun (2015, p.4), the Gurage area is divided into four agro-climatic zones, that is, alpine - 3200 meters above the sea level, temperate - 2300-3200 meters above the sea level, sub-tropical - 1500-2300 meters above the sea level, and tropical - 1100-1500 meters

above sea level. Alpine climatic zone is found around the Gurage Mountain, also called *Zebidar Mountain*, which has an elevation of 3,600 meters above the sea level (see Figure 2.2). The mountain divides the Gurage Zone into east and west; hence becoming a barrier to inter-group contacts between East and West Gurage people.



Figure 2.2: The Gurage Area, from Meyer (2014)

According to Menuta (2015, p.5), there is a mismatch between the geographical location and the traditional genetic nomenclature. For instance, Kistane is geographically located in the northeast part of the Gurage Zone while it is genetically classified under North Gurage. In the same manner, Silt'e is classified under East Gurage while geographically it is located in the southeast of the Gurage Zone (see Figure 2.2). Hence, some studies have made a distinction between the name of the geographical location and the nomenclature of the genetic classification of the languages. In Menuta (2015, p.5), east vs. west division was used for the geographical location alone, taking the Gurage Mountain as a reference point. Oriental Gurage and Occidental Gurage were used to express East and West Gurage languages respectively, a notion which had been adopted from Leslau (1951, p.212). In Figure 2.2, K'abeena and Libido are non-Semitic languages spoken in the Gurage area.

Many of the Gurage varieties have their own sub-varieties. For instance, Muhar has two sub-varieties the *Dessa* and the *Aklil*. Local peoples distinguish the two varieties as *anə-bet* and *ədi-bet* (Hetzron, 1977, p.5; Menuta, 2015, p.5). According

to Menuta (2015, p.7), Muher was originally called *Ohyə* literally an ‘udder’. The highland part of Muher was called *Nənn Ohyə* ‘the upper Ohye’ while the lowland part was called *Tət Ohyə* ‘the lower Ohye’. An oral story suggests that the *Ohye* people used to worship traditional deities, such as *Demwamwit* ‘god of fertility’ and *Bozhe* a ‘thunderbolt’. When Abune Zena Markos, a priest and missionary of the Ethiopian Orthodox faith, came to the Ohye and preached Christianity, the majority of the people converted into Christianity and abandoned their traditional beliefs and practices. Therefore, Abune Zena Markos called the people *muhir* or *mihur* ‘educated’. Therefore, Muher became the present designation of the people and their land (Hetzron, 1977, p.5; Menuta, 2015, p.7). Meyer [in discussion] disagrees with this oral explanation and argues that the name of the group in concern is *m<sup>w</sup>ohir* – which is not connected to the Ethiosemitic root *m-h-r* ‘learned, educated’.

According to Hetzron (1977, p.6) Chaha is derived from čəxja ‘a meeting plain land’. Gura is a variety of Chaha, and it is spoken around the village of *Aftir and Gura-Megnase*. Local people call it *Gurina*. Chaha and Gumer were traditionally called *əgir-angt čəxa* literally ‘neck and leg Chaha’, which actually means the ‘upper and lower parts of Chaha’ (Menuta, 2015, p.6). Gumer is also considered as a dialect of Chaha (Rose, 2007, p.403). Inor is very similar to *Ener* which is spoken in a area between Inor and Endegagn. Endegagn itself is closely related to Inor (Ahland, 2010, p.4; Hetzron, 1972, p.8; Hetzron, 1977, p.4). Mesqan is etymologically from *məsk’in* ‘cross’ (Menuta, 2015, p.7; Ousman, 2015, p.1). It is believed that Mesqan people were previously followers of Christianity. Most of them later converted into Islam in the 16th century during the Jihad of Ahmed Gagn (1506-1545). According to Menuta (2015, p.7) a significant number of Silt’e speakers live in and around Butajira (the town in the Mesqan area). I have also observed this during the field work. Kistane is derived from *kiristan* ‘Christian’. Kistane is also called Soddo. Kistane district is located in the northeast part of the Gurage Zone. Isolated pockets of Kistane speakers are also found in Oromiya Region, in the villages called Boleya, Gedamba, and Gewe at about 34 kilometres southwest of Addis Ababa (Menuta, 2015, p.7). According to Hetzron (1972, p.6), an isolated occurrence of Kistane is also found on the Gelila Island of the Lake Wonchi, near Ambo.

It has traditionally been believed that Silt'e speakers migrated from south Arabia. They first moved to Harar (east of Ethiopia), and then to the current location. Silt'e and Wolane speakers are believed to have the same ancestors. There has been an oral tradition which claims that the founding father of speakers of the two varieties separated at *Alaaba-kullito* (a place in the south part of Ethiopia) (Crass & Meyer, 2001, p.180). According to Crass & Meyer (2001, p.180), the name Silt'e was probably derived from *sultan* 'a sovereign governor'. In general, as can also be seen from Figure 2.2, the Gurage speakers are surrounded by the speakers of Cushitic and Omotic languages, and many of the Gurage varieties have a long history of contact with Cushitic and Omotic languages. Kistane has a strong contact with Oromo, Mesqan with Libido, Silt'e with Oromo, Alaaba and Hadiyya. Endegagn has also a close contact with Hadiyya. Similarly, Inor and Chaha have a strong contact with Yem. Muher has a contact with K'abeena, and Chaha with Yem.

### 2.2.2 Past Sociolinguistic Situation in the Gurage Area

There have been different oral traditions about the origin of the Gurage language and people. According to Taye (1954, p.48) and Meyer (2011, p.1221), the term 'Gurage' originated from *Gura* - a place in Eritrea. According to the oral tradition, the suffix *-ge* stands for 'land' or 'village' while the word 'Gura' means 'the people', further suggesting that 'Gurage' stands for the land or the country of the Gura people. There are other similar words that seem to substantiate the argument; for example, *Harar-ge* 'the land of the Harari people', *Abessh-ge* 'name of a district in Gurage Zone', *Arat-ge* 'a place in Kistane' (Menuta, 2015, p.4). According to Menuta (2015), 'Gura' is a common word among the Gurage communities which means 'left'- the opposite of 'right'. Hetzron (1972, p.7) argues against this oral tradition about the origin of the Gurage people and the language. He argues that there is no any linguistic evidence that substantiates the traditional claim.

The Gurage people were a politically unorganized community, and this political set-up has weakened the Gurage people (Menuta, 2015, p.8). A number of factors are often mentioned as the causes of this decentralization. One of the factors is the invasion of Ahmed Ibn Ibrahim al-Ghazi or Ahmed Gagn (1524-1543), a religious

warrior who invaded Gurage. The invasion instigated a number of conflicts among the Gurage community which later resulted in the formation of fragmented and decentralized several small communities. The second cause of decentralization is the political division and hostility during 'Zemene Mesafint' (1769-1855) in the history of Ethiopia. This era is characterized by a complete absence of social order, and lack of a central government. Among the Gurages, there were numerous war lords who were fighting each other for power, land and other resources. For instance, there were wars between Chaha and Inor, Ezha and Chaha, Muher and Ezha, Mesqan and Kistane, Mesqan and Silt'e and others (Dinberu et al., 1995, p.217).

After the end of Zemene Mesafint, by the end of 19th century, there were also wars among Christian and Muslim Gurage communities. These wars were considered as another factor for the decentralization of the Gurage community (Worku, 2005, p.934). The division among the Gurage community continued in the 20th century, partly due to an exclusive approach of the administrative arrangement. For instance, during the Emperor Haileselesie era (1930-1974), different areas of Gurage such as Kistane and Mesqan were included in *Hayk'och-and-Buttajira Zone* together with other non-Semitic language speakers such as Oromo. Moreover, the speakers of Sebat Bet Gurage, together with Oromo, were categorized under *Chebo-and-Gurage Zone*. Endegagn was included in the Kambata Zone (Worku, 2005, p.930). Similar administrative arrangement was used during the communist Dergue regime (1974-1991). Some parts of Oromo (Weliso) was merged with Gurage. According to Menuta (2015, p.9), this administrative division throughout the history hindered the formation of a strong Gurage administrative unit which could have created a strong bond among Gurage communities. After 1993, the ethnic federalism has been implemented in Ethiopia. Even earlier at this time, Gurage communities were initially merged with speakers of Cushitic (Hadiyya) and of Omotic (Kembata) (Aalen, 2000, p.1411). Later, however, the Gurage people were grouped under the Gurage Zone within South National and Nationalities People Regional State.

Since then, the Gurage communities have been institutionalized together, with its own centralized administration. This was a good opportunity to unify the Gurage community and the disintegrated language and culture (Menuta, 2015, p.9).

Nonetheless, re-organizing the Gurage ethnic group and standardizing the language varieties have still sustained as challenges. The biggest challenge has been ethnic diversity among Gurages, and disagreement among the Gurage ethnic groups and the political elites on the linguistic materials to be considered during the standardization process (Meyer, 2018, p.5-8; Menuta, 2015, p.8). I also observed the same situation during a discussion with an official at the Gurage Cultural Bureau in the Gurage Zone. In the case of Silt'e, there was a great ambition among the Silt'e elites to standardize the language and to use for schooling and administration. Nonetheless, since the independence in 2001, the standardization issue has been largely abandoned (Meyer, 2018). Though Silt'e has been taught in the elementary schools, Amharic has remained the main administrative language.

### 2.2.3 Current Demographic and Sociolinguistic Issues

According to the current federal structure in Ethiopia, Gurage people live in the Gurage and Silt'e zones. Figure 2.3 shows the administrative map of Gurage. Administratively, the Gurage Zone is divided into thirteen districts: Abeshge, Chaha, Endegagn, Inor, Ezha, Gyetu, Gumer, K'abeena, Libido, Mesqan, Kistane, Muhir-Aklil and Wolane. Silt'e was the fourteenth district in the Gurage Zone until it became an independent zone in 2001. The Gurage administrative area is demarcated by Oromiya region in the north and northwest and southwest, and Yem Special District in the southwest and Hadiyya Zone in the south.

Not all Gurage people live in the Administrative Zone. The Gurage people are very mobile (Henry, 2001, p.6; Menuta, 2015, p.16). More than half a million Gurage people live outside the Gurage land. The current Gurage settlement is the outcome of several historical and social waves; the dispersion of the Gurage people across the Gurage land, and the present day's Gurage community settlement is the outcome of various social dynamics. According to Dinberu et al. (1995, p.82), movements within Gurage were largely from east (Kistane) to the west. These movements were motivated by internal (political, social and economic situations within Gurage) and external factors. Pressure from the dominant neighboring languages is the main external factor. For instance, the movement of Kistane from east to west towards



Mesqan, Muher and Ezha, following the chain of the Gurage Mountain (Menuta, 2015, p.15) was probably the result of the pressure from Cushitic languages mainly from the Oromo. According to Menuta (2015, p.15), it is because of these movements that today there is *ädi* 'I' group (a Kistane community) in Muher though Kistane is geographically not adjacent with Muher. According to Dinberu et al. (1995, p.82), the Gurage area was also influenced by the movement of Highland East Cushitic language speakers, mainly Alaaba speakers who occupied the land of Chaha and Gumer and advanced further to Inor.

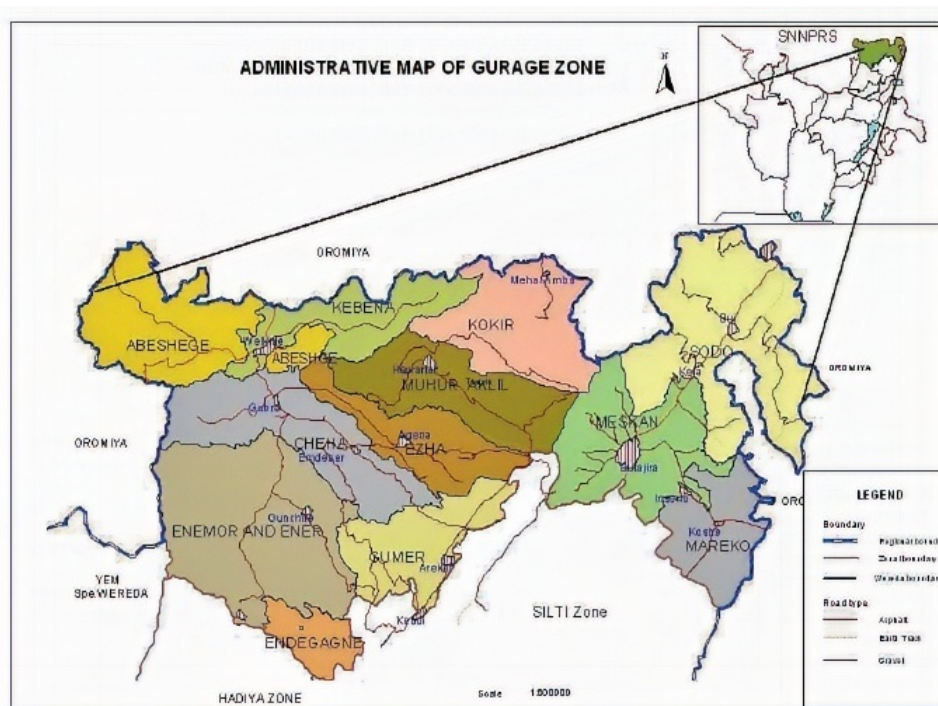


Figure 2.3: Administrative map of Gurage, from Menuta (2015)

Crass & Meyer (2001, p.179-180) also indicate that the Silt'e people moved from east of Ethiopia to their present day's position in the Silt'e Zone and Wolane districts, then to the Chaha district. Later, after the expansion of the Silt'e people, according to Menuta (2015, p.15), there was a contrary west to east movement across the Gurage land which influenced the settlement of the Silt'e people. For instance, Chaha speakers moved from lowland to highland and formed a Gumer community. Then, the Gumer community advanced southeast and occupied some of the lands which were previously occupied by Silt'e. Hence, according to Menuta (2015, p.15), place names such as *Chacho*, *Medercho*, *Wenabo*, *Sente* and *Aselecha* previously be-

long to Silt'e. Nowadays, however, they are parts of Chaha. Crass & Meyer (2001) also seems to suggest that Mesqan speakers came from Chaha via Gumer which shows another west to east movement.

Out-movement, the movement of the Gurage people from the Gurage area to other parts of Ethiopia, has also its own influence on the current linguistic situation of the Gurage area. As indicated in 2.2.1, Gurage communities are highly mobile. According to Nida (2000, p.43), out-migration is called *fannonət* among the Gurage communities. Nida (2000, p.43-55), discusses three causes of Gurage emigration: occupation by Menelik II, Italian occupation and the construction of Gurage roads. In 1888, the Gurage communities were incorporated into the broader part of Ethiopia by Menelik II. Then, the shortage of land combined with tribute and taxes imposed by the then feudal administration forced some of the Gurages to emigrate to the urban area, mainly to Addis Ababa. Moreover, following the Italian occupation in 1935, *merkato* 'a labor market center' was established at the center of Addis Ababa. The labor market which created a wedge, attracted the attention of some of the Gurage youth. The construction of roads and schools and the awareness created following these changes have also accelerated the migration of the Gurage communities (Baker & Pedersen, 1992, p.127-128; Zewde, 2002, p.23). Nowadays, many Gurages are businessmen in the big cities in Ethiopia where Amharic is the dominant language; Amharic has been imported to Gurage area partly by these traders. As a result, today, Amharic is used as language of administration, medium of instruction and language of elites in the Gurage area. The Gurage varieties are used only as spoken languages (except Silt'e), and they are limited to home and market communication. Henry (2001, p.6) suggests population density in the Gurage area as another cause of the out-movement migration.

Out-movement migration is just one of the several ways through which Amharic was imported to the Gurage area. Amharic has become a dominant language because of other two reasons: national enforcement and forced settlement (Menuta, 2015). As Amharic was the national language of the country (today Amharic is recognized just as the language of the Federal Government), it was imposed on all nationalities in the country, including the Gurage communities. Indeed, as the lan-

guage of the Federal Government, it is still taught in all elementary and secondary schools throughout the country. There was also a nationwide resettlement campaign during the Derg regime, from 1973-1986. During this forced settlement, some Amharic speakers moved to the Gurage area.

From social point of view, different Gurage groups are inter-related through marriage. Often, in Gurage, marriages take place among people that do not belong to the same kinship. The most preferred marriage is between distant ethnic groups within Gurage, but with comparable social and economic statuses. Monogamy is a common practice, but polygamy is also practiced especially among Muslim Gurage communities such as Mesqan and Silt'e (Menuta, 2015, p.12). According to Henry (2001, p.6), during intermarriage and migration, men maintain their original clan affiliation. For example, if a Chaha man marries a Mesqan woman or migrates to the Mesqan area, his family maintains Chaha identity. Wives usually adopt the clan of their husbands. There are also different ways in which the Gurage communities socially organize themselves. Ethnic hierarchy is one means of organization which consists of *Tib*, *Bet*, *Wefencha* and *Den* (Menuta, 2015, p.9). *Tib* is the highest social hierarchy. It encompasses almost all Gurage clans. *Bet* is the second broad social hierarchy which also consists of several Gurage clans. *Wefencha* is the third social hierarchy which consists of sub-clans within the clan. *Den* is the fourth and the smallest social hierarchy which usually refers to members of a sub-clan (Menuta, 2015, p.8). There are other various social structure including *Iqub*, *Idir* (traditional saving system), and labor associations. Labor associations can be different form: *Weje* (herding cattle by takings turns), *Gyez* (traditional cooperative farming and mat making), and *Wusacha* (scraping of ensete plant in group) (Menuta, 2015, p.9-10).

Various social groups exist among Gurage such as *Seven House Gurage*- a tribal confederation consisting of Ezha, Chaha-Gumer, Muher, Meqorqor, Inor, Gyeto and Endegagn (Nida, 1996, p.134; Rose, 2007, p.405). According to Henry (2001, p.7), there was Amist-bet Gurage 'Five House Gurage' before 1875, consisting Chaha, Muher, Endegagn, Gyeto and Ezha. The Sebat Bet Gurage came into existence after Muher Aklil and Wolane (Wolane is not considered Sebat Bet Gurage today)

joined Amist Bet Gurage in 1889. Another institution is *Seven House of Gogot* (Drewes, 1996, p.72). This includes Mesqan, Silt'e, Wuriro (consisting Allicco and Abbecco), Azarnat and Barbare, Dalocca, Hulbarag and Alaaba. This group consists of mixture of languages - some of them non-Semitic languages, for example Alaaba. Gogot and Kistane are not members of any of these two social groups.

Shack (1966) and Zewde (2002) also discuss different forms of traditional legal systems. For instance, according to Zewde (2002, p.20-21), there are traditional systems of administration in Gurage communities which manifest themselves in three level of authority: village, clan and region. These legal systems declare different sets of laws such as *Sabugnat*, *Agar* and *Yajoka and Gordanna Sera*. Sabugnat is the lowest organization consisting of 10-60 households. Sabugnat exercises social, political and judiciary functions. Its main responsibilities include providing gifts during funerary and wedding for first-born son, ensuring equitable use of grass lands, and arbitrating in cases of dispute. In the case of dispute, it administers sanction that ranges from assigning a number of guests to be dinned and wined at the expense of the culprit (Yekka) to social exclusion. Agar performs the roles of Sabugnat at higher level, on estimated number that ranges from 1,000 to 6,000. It administers churches and communal lands, following the rules laid down by the Gordanna Sera. It also administers sanctions which range from Yekka to social exclusion. The highest level of sanction is the ultimate curse 'May the eye of Agar eat you', the form of evoking the total dishonor of the community. Yajoka and Gordanna is the highest administrative hierarchy. Its main responsibilities include defending the community from external aggression, enacting administrative and judiciary laws and supervising the implementations of the fundamental rules of law. According to Henry (2001, p.6), Gurage people never had a centralized political leadership; instead they had a segmented political system with authorities vested in clan heads and elders.

With regard to the economic activities, according to Menuta (2015, p.11) Tilahun (2015, p.10) and Baker & Pedersen (1992, p.126) , the Gurage people are predominantly farmers. The main crops of the area include wild banana 'Ensete Ventricosum' from which they make different kinds of food, mainly a kind of bread called *wusa*. Wild bananas usually grow in the lowland and semi-highland areas of the

Gurage Zone. The extreme highlanders, such as Muher do not plant wild bananas since they are not productive in such a cold area. Besides wild bananas, the Gurage people produce different kinds of cereals such as barley, wheat, maize, pepper, lentils, 'teff' (*Eragrostis tef*) and commercial crops such as coffee. Gurage people also rear animals such as cows, sheep, goats, horses, mules and donkeys. Above all, they are well-known as traders even among the remaining Ethiopian communities. The Gurage people are either Christian or Muslims. Muher and Kistane are predominantly Orthodox Christians. Besides the two religions, there have existed practices of traditional religions. Among several gods and spirits worshiped in Gurage, one is *Waq* 'God of wealth' which is also called *Ogyet* in Chaha, *Mando* in Gyetu, *Ingyeber* in Ezha, *Yeber* in Inor and Ener, *Yemwarer/Yesu* in Muher. There are also gods called *Demwamit* and *Bozhe* 'thunderbolt' (Menuta, 2015, p.7).

With regard to the population size, according to ENSA (2007, p.6-8), the estimated Gurage speakers number 1,125,929. This figure includes Gurage speakers who live outside the Gurage Zone. The population size of each district in the Gurage Zone is presented in Table 2.2. This is not the exact number of the native speakers since the Ethiopian National Statistical Agency Report does not provide the native speakers of each Gurage variety. However, it provides the number of residents of each district which its division is mainly based on the ethnic background of residents. Hence, the numbers of residents of the ethnically defined districts (e.g., Ezha Wereda, total population 84,905) was considered as the numbers of speakers of the varieties of Ezha, in this particular example. Therefore, it is also important to note that the figures in the table do not exclusively show the native Gurage speakers since there are also some non-Gurage speakers who live in the Gurage Zone. In total around 1,280,483 people live in the Gurage Zone. The estimated number of Gurage speakers in the Zone is 1,030,621 (note that this figure stands only for those who live in the Gurage Zone) (Samia, 2007, p.201). There are 750,398 residents in the Silt'e Zone. The estimated native Sil'te speakers (including those who live outside the Silt'e Zone) are 801,091.

Table 2.2: Population size in nine districts in Gurage area

Districts	Urban		Rural		Total
	Male	Female	Male	Female	
Chaha	4,574	4,418	52,277	54,682	115,951
Endegagn	395	371	22,459	25,946	49,171
Ezha	1,8202	1,890	38,441	42,754	84,905
Gumer	1,412	1,511	36,083	41,172	80,178
Inor	3,380	3,210	75,871	85,309	167,770
Mesqan	6,1142	5,274	70,282	74,112	155,782
Muher	401	322	40,621	46,412	87,756
Kistane	6,895	6,825	60,235	60,728	134,683
Silt'e	24,757	22,340	339,351	363,950	750,398

A survey conducted by Menuta (2015) shows that almost all Gurage speakers are bilinguals. Speakers of the Gurage varieties have a frequent contact among each other. Many of them have contact with Chaha, compared to other Gurage varieties. The majority of the participants believe that they understand Chaha better than other Gurage varieties. They also attribute positive value to Chaha. The most common place of contact include market places, in towns and Yejoka - a traditional courting place. Besides, according to this survey, people who are geographically adjacent have more contact than remote people, for example, Mesqan speakers have frequent contact with Kistane. Gurage Mountain was mentioned as the main geographical barrier to contact among the Gurage speakers. For instance, Mesqan and Kistane do not have a frequent contact with the speakers of Chaha, Muher and Inor because of this barrier. Lack of public transport and hostility among the Gurage clans are indicated as other obstacles. In market places, Gurage is more frequently used than Amharic. In schools, religious places and administration, Amharic is the dominant language. According to this study, Amharic (63%) is more frequently used at home than Gurage (42%).

---

## Chapter 3

# Ethiosemitic Languages

This unit discusses the Ethiosemitic languages with a particular focus on the controversies related to the origin of the languages, the main classification proposals that have been made and the limitations which have been observed in the classification proposals. The chapter contains three sections. The first section presents various hypotheses about the origin of Ethiosemitic languages in general and the Gurage varieties in particular. The second section discusses the previous classification attempts of Ethiosemitic languages. The third section highlights the challenges that linguists have faced in the attempt of classifying Ethiosemitic languages. The majority of the studies reviewed in this chapter employed morpho-syntactic features with the aim to classify Ethiosemitic languages genealogically. The author recognizes that the approaches used in these studies are fundamentally different from the assumptions of the areal classifications aspired by the present study. However, the author also believes that the genealogical classification attempts and their limitations can undoubtedly be essential inputs for the areal classification anticipated by the present study.

### 3.1 The Origin of Ethiosemitic Languages

The origin of Ethiosemitic languages has been a subject of debate for decades among Semitists. The debate emerges from the enduring lack of clarity about the origin of the Semitic languages in general. There are two hypotheses with regard to the

origin of the Semitic languages: traditional theory and the Africanist view (Demeke, 2001, p.59; Menuta, 2015, p.10).

The traditional view assumes that some speakers of Afroasiatic languages initially migrated from Africa to Asia. Through time, after their separation from the rest of Afroasiatic speakers, their language developed its own distinct features in Asia and formed a proto-Semitic language. Then, the speakers of one of the affiliates of the proto-Semitic language migrated from South Arabia to Ethiopia (Ehret, 1988, p.641, Fleming, 1668, p.354, Gragg, p.163; Hetzron, p.15-19 Hetzron, p.16). Based on this assumption, today, many scholars believe that Ethiosemitic languages descended from the Proto-Semitic language family, particularly, from the South Semitic branch which was spoken in the South Arabia. According to this theory, these migrants first moved from south Arabia to north Ethiopia. Then later, some Semitic speakers moved from the north to the south (Hetzron, 1972, p.36). Those who took the direct route from north to the south are the speakers of Gunnən Gurage and Gafat which form together Outer South Ethiopic (see Hetzron, 1972, p.36). However, those who moved from north Ethiopia to east Ethiopia and then moved to south constitutes the speakers of Amharic, Argoba, Harari and East Gurage (Silt'e, Zay and Wolane) which form together Transversal South Ethiopic (Hetzron, 1972, p.36).

The alternative hypothesis, contrary to the first one, is that the origin of Ethiosemitic languages is Africa. Therefore, it is called Africanist view. According to this proposal, Ethiosemitic is a descendant of the Afroasiatic language which had been spoken in Africa in pre-Semitic era. The hypothesis implies that Ethiosemitic languages had been spoken in Africa before the expansion of the Semitic languages across Asia and north Africa. According to Demeke (2001, p.59-60), two explanations are often provided to support this proposal. First, among six of the Afroasiatic sub-families (Semitic, Cushitic, Omotic, Berber, Chadic and old Egyptian), only a few of them are spoken in Asia while all of them are spoken in Africa. Among these, three of them are spoken in Ethiopia (Semitic, Cushitic and Omotic). Based on this, Ethiosemitic is assumed as the mother language of all Semitic languages.

Another reason is that more Semitic languages are spoken in Ethiopia (around 16) than in Asia. Hence, based on the 'least move principle' they assume that the



source of all Semitic languages is Africa, particularly Ethiopia (see Hudson, 2000). Though some scholars, for example Demeke (2001) and Hudson (2000), believe that the second proposal is more convincing than the first one, the current archaeological and linguistic evidence largely suggests the contrary. In the support of this, recent phylogenetic studies suggests that Ethiosemitic languages are descendants of South Semitic - a presumptive branch of Semitic language (cf.: Kitchen et al., 2009). However, it is too early to reject the claim that the origin of all Semitic languages is Ethiopia. Rigorous comparison of Ethiosemitic languages with other Semitic languages and combining this with recent archaeological findings may settle the debate sometime in the future. It is up to the future to decide!

There are three hypotheses with regard to the origin of Gurage varieties. The first one assumes that the Gurage people came from present day Eritrea, an area called *Akale Guzay* during the reign of Amde-tSION (1314-1344). The ancestors of the Gurage speakers, therefore, assumed to be the soldiers of King Amde-tSION. This hypothesis has been supported by oral tradition (see §2.2.2). Some Gurage speakers even today believe that their origin is from a place called *Gura*. The second hypothesis is that the ancestors of the Gurage speakers migrated from east of Ethiopia during the expansion of Ahmed Gragn (1524-1543), a Muslim warrior. During the expansion, the Islamist militia and troops immigrated to the Gurage area. The third hypothesis is that the Gurage varieties are originally African; they were spoken in the south part of Ethiopia even before the reign of Amade-tSION and the expansion of Ahmed Grange. As argued by Menuta (2015, p.13), the idea that the Ethiopian Semitic is the source of all the Semitic languages also suggest that Gurage varieties were probably originally Ethiopian, and they were spoken in Ethiopia from the beginning.

Menuta (2015, p.13) also argues that a number of features that are present in the Gurage varieties are absent in north Semitic languages. However, there is no convincing evidence whether these features are innovations within Gurages themselves or borrowings from neighboring Cushitic languages such as Oromo, Hadiyya, Alaaba and Omotic languages such as Yemsa (see Figure 2.2). According to Hudson (1977, p.129) several Proto-Ethiosemitic features have been identified in Gurage.

These features include main verb markers (e.g., Leslau, 1967, p.122-125 and Hetzron, 1968), and the archaic vowel of jussive (Leslau, 1982). Based on these features, Hudson seems to imply that the features are not innovations within Gurage. Menuta (2015, p.13-14) further argues that the archaic jussive vowels, the main verb markers, palatalized velar consonant phonemes: /k<sup>j</sup>, g<sup>j</sup>, k'<sup>j</sup>, x<sup>j</sup>/ etc. are typical Gurage features, but he did not make a claim whether they are archaic Proto-Semitic features or innovations within Gurage. However, Menuta (2015, p.13-14) recognizes the influence of Cushitic languages especially on the East Gurage languages in terms of, for example, lexicon and phonology. Both the origins of the Gurage varieties and the Gurage people have largely remained the matter of speculation.

Cognate frequency was also employed to determine the time of separation of Ethiosemitic languages from their ancestors. Fleming (1968) is worth mentioning. The study employed Glottochronology to estimate the origin and the genetic relationship among Ethiosemitic language, Modern South Arabia and Old Epigraphic South Arabia languages. This method employs the changes in the basic vocabulary to hypothesize a period of divergence among genetically related languages. Fleming (1968) used the Swadesh list of 100 basic words. The Genetic branching among the languages was estimated under the assumption that 80-85% of basic vocabulary is retained in every language for 1000 years. In other words, 15% to 20% 'dissociation' occurs among 'basic concepts and basic vocabularies' within 1000 years. Based on this assumption, Fleming (1968, p.363) estimated that the South Ethiosemitic separated from North Ethiosemitic between 700BC to 300 BC. However, the basic vocabularies of only a few languages were compared: Tigre from North Ethiosemitic and Chaha and Wolane from the South. He estimated that the separation between the North and South Ethiosemitic occurred either in South Arabia or around the Red Sea.

According to his estimate, the separation between Ethiosemitic and Modern South Arabia took place somewhere between 3300BC and 2600 BC. Using the same method, Fleming (1968, p.363) estimated that the diversification of the South Ethiosemitic languages began between 300 BC and 100AD. However, the estimation was made based on the basic vocabulary of three languages: Chaha, Amharic

and Harari. While Fleming indicated that Tigre, Tigrigna and Ge'ez may share the same ancestor, his estimation was not based on detailed investigation of the South Ethiosemitic languages. He also seems to suggest that South Ethiosemitic languages probably derived from South Arabia not from Ge'ez or other proto-language in the North Ethiopia. Bender (1966) also employed Glottochronology to estimate the date of separation of other Ethiopsemitic languages from Ge'ez. He assumed that Ge'ez is the proto-language of all Ethioemitic languages, the assumption which is today widely refuted (see Palmer, 1958, p.120). He examined the basic vocabulary of nine languages: Ge'ez, Tigrigna, Tigre, Amharic, Harari, Kistane, Chaha, Gyeto and Mesqan. According to his estimation, North and South Ethiosemitic languages separated around 2000 years ago.

### **3.2 The Classification of Ethiosemitic Languages**

Like the origin of Ethiosemitic, the internal classification of Ethiosemitic languages is debatable. Though many Semitists believe that Ethiosemitic languages descended from the South Semitic (see 3.1), they do not provide a completely convincing explanation about the historical development and the history of expansion of the speakers of the Ethiosemitic languages. Some believe that initially the speakers of Ge'ez, the old language of liturgy of Ethiopian Orthodox Church, migrated from South Arabia to Ethiopia, and the rest of the Ethiosemitic languages then probably descended from Ge'ez (Felhnan, 1996, p.205-206). However, this argument has recently been widely refuted, and the more plausible perspective available today is that Ge'ez is not the mother language of any of the Ethiosemitic languages (Demeke, 2001, p.64; Faber, 1997, p.6-7; Ullendorff, 1955, p.11-12). It is just a sister language of the other two North Ethiosemitic languages (Tigrigna and Tigre). According to Fleming (1968, p.358) and Goldenberg (1977, p.461), the disagreements in the classification attempts partly emerge from heavy inter-borrowing among several Afroasiatic languages and from the lack of adequate data on some of the languages.

Regardless of some disagreements (see Hudson, 2013; Voigt, 2009), many of the studies recognize that Ethiosemitic languages are broadly classified into North and

South Ethiosemitic. Though this classification has frequently been accepted by the experts of Ethiosemitic languages, the position of the South Ethiosemitic within the Ethiosemitic language family is controversial. Some linguists believe that South Ethiosemitic is a branch of the proto-Ethiosemitic language (Demeke, 2001; Hetzron, 1972). According to these scholars, both the North and the South Ethiosemitic languages descended from the same proto-family, i.e., Ethiosemitic. Nevertheless, there are also linguists who believe that South Ethiosemitic languages are direct descendant of the South Semitic (Fleming, 1968, p.363; Hetzron & Bender, 1976, p.4). The controversies often emerge from the profound grammatical, morphological and phonological differences between the South and the North Ethiosemitic languages (Demeke, 2001, p.70-71; Hetzron, 1972, p.22-28; Voigt, 2009, p.1375-1376). Glottochronology studies also suggest a direct branching of South Ethiosemitic from the South Semitic (Fleming, 1968, p.365).

According to Fleming (1968, p.358), it is difficult, however, to precisely claim which of the hypothesis is right since many of the South Ethiosemitic languages and South Semitic languages are not well studied. The evidence that is provided to justify the link between South Arabia and Ethiosemitic languages is inadequate. Hetzron (1972, p.15-16) provided three explanations to illustrate the link. (1) The presence of negation marking prefix *-al* in both branches. (2) A very close similarity in the second person singular masculine suffix *-k* in Ethiosemitic, and *-t* in South Arabian languages (*t - k*). (3) The presence of *yV-CCVC* jussive pattern in both branches, *yi-k'bər* in Ge'ez and *l-ik'bər* in Soqotri.

The internal classification of the South Ethiosemitic languages is less clearer than the classification of the North branch. This is because of the complex areal diffusion as presented in 2.1. Partly because of this intermingling, some previous works (e.g. Demeke, 2001 and Hetzron, 1972) have proposed classifications which reflect certain degree of differences. These differences are mainly due to the lack of adequate data on some of the languages, especially on the Gurage languages (Demeke, 2001, p.67; Gutt, 1980, p. 78; Hudson, 2000, p.84). Among the classification attempts that have been made so far, Hetzron (1972) and Demeke (2001) are among the most comprehensive ones; they dealt with almost all major Ethiosemitic languages. They have

provided more or less comprehensive classifications which included the majority of the Ethiosemitic languages. Many linguists believe that Hetzron (1972) is so far the most detailed and complete classification (Demeke, 2001, p.68; Goldenberg, 1977, p.461; Hudson, 2000, p.75; Rubin, 2008, p.92). Rubin (2008, p.92) also indicates that, after Hetzron (1972), there were no detailed classifications of Ethiosemitic languages partly because of the political instability in Ethiopia and partly since the work of Hetzron (1972) was complete.

Hetzron (1972) has been admired especially for its methodological selection and depth (Goldenberg, 1977, p.461; Kogan, 2005, p. 370; Palmer, 1978, p.584-585; Rubin, 2008, p.79) though Robert Leslau and Marcel Cohen also contributed a lot to the study of Ethiosemitic languages. The classification of Hetzron (1972), nevertheless, has also been criticized for its limitations in particular for the classification of Gurage languages (see Demeke, 2001; Goldenberg, 1977; Hudson, 2000; Rubin, 2008). The lack of data on some languages impeded Hetzron's work. Goldenberg (1977) also reviewed Hetzron (1972) and severely criticized the validity and the consistency of its classification parameters.

As a response to these limitations, Demeke (2001) recently reviewed Hetzron (1972) and the subsequent publication, Hetzron (1977) which exclusively focused on the classification of Gurage varieties - *Gunnən Gurage* and re-examined the classification of Ethiosemitic languages. He also considered the views of the critics of Hetzron's classification. As far as I know, Demeke (2001) is the most recent attempt towards re-examining the whole classification of Ethiosemitic languages predominantly using morphological parameters similar to Hetzron (1972). As a result, in the present study, Hetzron (1972) and Demeke (2001) are taken as the main sources to illustrate the classifications of the Ethiosemitic languages and to demonstrate the persisting limitations in the classifications. The classifications that have been proposed in other studies are given as supplementary evidence to strengthen the arguments provided in the classifications of the two scholars. I will first present the classification of Hetzron (1972), and then introduce the alternative classification proposed by Demeke (2001).

Hetzron (1972) classified Ethiosemitic languages into North and South Ethiosemitic.

The classification of Ethiosemitic languages into the North and the South is relatively well established. According to Demeke (2001) and Hetzron (1972), the classification of the two groups is mainly based on six language features: (1) perfective and imperfective aspectual contrast in the penultimate consonants of type <sup>2</sup> ‘A’ verb (perfective: *səbərə* ‘broke’ in the North and *səbbərə* ‘broke’ in the South; imperfective: *yī-səbbir* ‘breaks’ in the North and *yī-səbir* ‘breaks’ in the South), and the presence of palatal vowel in type ‘B’ verb in the North but not in the South (*fəśśəmə* in Ge’ez-North) and (*fit’t’əmə* in Kistane - South). (2) Broken plurals in the North Ethiosemitic, but concatenate morphemes in the South. (3) High degree of number of vocabulary similarity (cognates) among the North Ethiosemitic languages (Tigre, Tigrigna and Ge’ez); the three North Ethiosemitic languages share more cognates among each other than the South Ethiosemitic languages. (4) Negation morphemes: *i-* in Ge’ez and *ay-* in Tigrigna, but *al-/l-* in the South Ethiosemitic languages as in *al-hedəm* ‘he did not go’ in Amharic. (5) The cardinal number ‘nine’ maintained the Semitic root *tš* in the North, *tšatte* in Tigrigna, but adopted the Cushitic root *-zht’ñ* in the South, *zət’t’əñ* in Amharic and *ziht’əñ* in Harari. (6) Lack of geminate consonants in type ‘C’ verb in the North, but geminates in the South; for examples, *yī-barix* ‘he blesses’ in Tigrigna, *barik* ‘he blesses’ in Tigre, but *yī-barrik* ‘he blesses’ in Amharic.

---

<sup>2</sup>See Table 2.1 for the distinction between Type ‘A’, ‘B’ and ‘C’ verbs

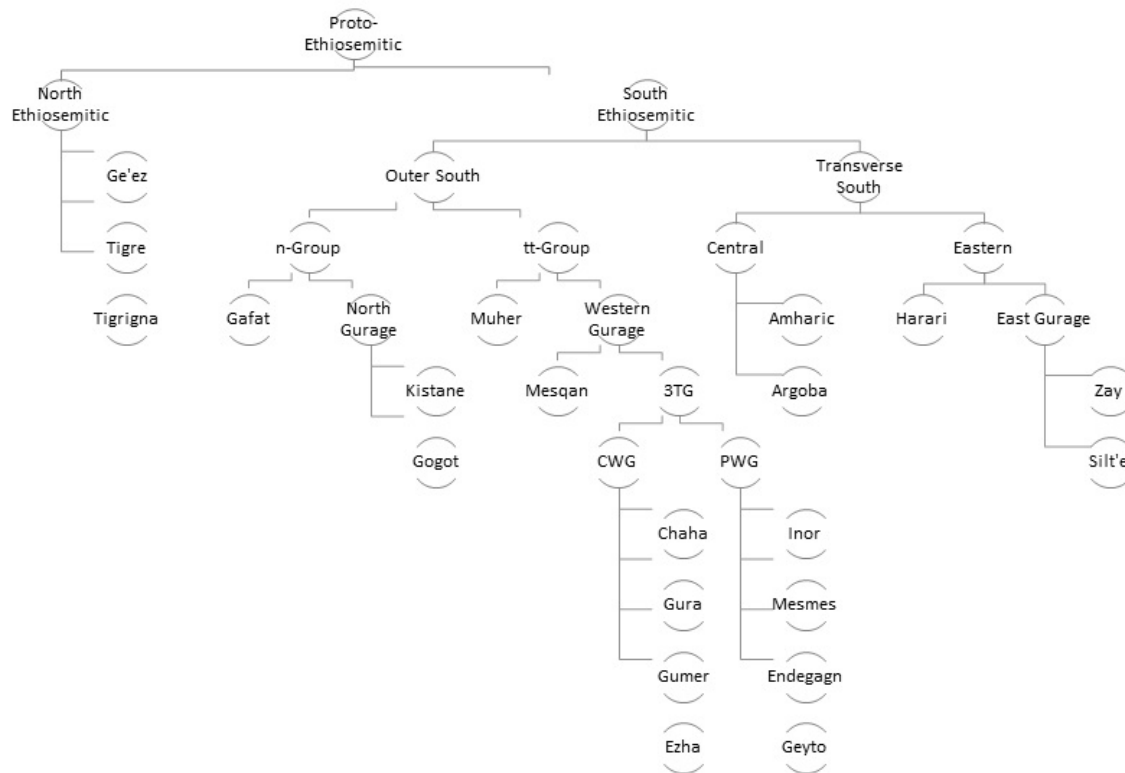


Figure 3.1: Classification of Ethiosemitic languages according to Hetzron (1972)

Hetzron (1972) classified the South Ethiosemitic languages into Transverse South and Outer South Ethiosemitic (see Figure 3.1). This division was based on three features- *isoglosses* in the words of Hetzron. (1) The root of the word ‘to laugh’ is *shk’* in Transverse South as in *sak’ə* (Amharic) and *səhak’ə* (Argoba), but the Outer South Ethiosemitic developed another root which begins with an <sup>3</sup> emphatic consonant, e.g., *sak’ə* (Gafat) and *ḍak’ə* (Gunian Gurage). (2) Transverse South Ethiosemitic lost gender distinction in feminine and masculine while Outer South Ethiosemitic preserved the distinction. (3) Main Verb Marking: Semitic languages, in general, make a distinction between main indicative clauses and the subordinate ones. In Transverse South Ethiosemitic (both in Eastern and Central Transverse), the existential auxiliary *-all* is added to the verb in the main clauses. In the following Amharic examples, the auxiliary marking *-all* is used as a marker of the main clause

<sup>3</sup>Emphatic consonants are pharyngealized obstruents.

(3a), but it is absent in the subordinate clause (3b).

3. (a) *issu məsəhaf-u-n šifw-all*  
 he book-DEF-ACC write.IFV-AUX.3.M.SG.  
 ‘He has written the book.’

- (b) *issu məsəhaf-u-n kə-šaf-ə*  
 he book -DEF -ACC if-write.PVF-3.M.SG  
 ‘If he writes the book.’

However, Outer South Ethiosemitic languages use special suffixes to distinguish the main clause from the subordinate ones, and these suffixes are called *Main Verb Markers*. According to Hetzron (1972, p.38), the Proto-Semitic main verb marking is *-u* after words that end with consonants and *-n, na or ni* after words that end with long vowels. Outer South Ethiosemitic languages use almost the Proto-Semitic suffix combined with a suffix *-t* for the main verb marking (e.g., Gogot *əsəbr-u* ‘I break’, *tisibri-n* ‘you (F.SG) break’; Muher: *tisibri-tt* ‘you (F.SG) break’). The Transverse South consists of two branches: Central Transverse and Eastern Transverse. The Central branch consists of two languages: Amharic and Argoba. These two languages are characterized by (1) complete conservation of geminates in type B verbs as illustrated below.

4. (a) **Amharic:** *nəggədə* ‘traded’, *yi-nəggid* ‘trades’, *yi-nəggid* ‘shall trade’  
 (b) **Argoba:** *nəggədə* ‘traded’, *yi-nəggid* ‘trades’, *yi-nəggid* ‘shall trade’

According to Hetzron (1972, p.40), an innovation of Argoba is that the gemination is extended to type C jussive which is not the case in Amharic: *yi-gallib* ‘shall ride’ in Argoba and *yi-galib* ‘shall ride’ in Amharic. (2) Another shared feature is a very close similarity between the two languages in terms of present perfect form, which is composed of converbs and auxiliary with the root *hwl*. In Amharic, *-all* serves as an auxiliary. The same form of the auxiliary is used in Argoba except for slight phonological changes and addition of *-d* to the Argoba auxiliary: Amharic, *səbr-ən-all* ‘we have broken’; Argoba, *səbri-d-ən-əl* ‘we have broken’. 3) 3rd person



possessive pronoun suffixes are *-u/-wa* in both Amharic and Argoba contrary to the Eastern Transverse languages, for example, *zo/-ze* in East Gurage. (4) In both languages, the relative particle *yə-* is used before the perfect and *yam-* or *imm-* before imperfect; for example in Amharic, *yə-mət't'a* ‘who came’ and *yəmmi-mət'a* ‘who comes/will come’. According to Hetzron (1972, p.41), no other language has these suffixes. Furthermore, (5) only these two Ethiosemitic languages have introduced *n* in the numeral ‘one’: *and* in Amharic, and *hand* in Argoba.

Hetzron (1972) further classified Eastern Transverse South into Harari (spoken mainly in the city of Harar) and East Gurage which also consists of Zay and Silt'e. There are four features that Eastern Transverse South language share, according to Hetzron (1972, p.42-44). (1) Harari and Eastern Gurage use compound verbs (main verb plus main clause marking auxiliary) in <sup>4</sup>relative clauses. For example, *yi-səbrizaal* ‘who breaks’ in Harari, *yi-səbran* ‘who breaks’ in Silit'e (cf.: Amharic *yəmmi-səbr*). Amharic does not admit compound imperfect in relative function. (2) Palatal quality of type B verbs: if a consonant preceding the high front vowel /e/ is palatalizable, the consonant undergoes palatalization and /e/ is dissolved by the consonants and shorten to ə: \**yizebbir* - *ižəbbir* ‘he breaks’. Otherwise, just /e/ is reduced to /ə/. This process is extended to jussive in Harari and East Gurage: for example, *yišəmk'i* ‘shall break’. (3) These languages are also characterized by loss of germination in the second radical in Type B verbs (cf: Amharic: *yi-rəkkəb* ‘receives’; Harari: *yi-rəxəkbi* ‘receives’).

According to Hetzron (1972), the Outer South Ethiosemitic branch consists of two sub-branches: n-group and tt-group. This classification is based on the main verb markers (*-tt* and *-n*) as in Kistane and Dobbi: *yisabrimu-n* and in Muher *yisabrimu-tt*, ‘they will break’. The n-group in turn consists of Gafat (a dead language) and another branch - North Gurage which consists of Kistane and Gogot. According to Hetzron (1972, p.57-58), there is little evidence for genetic relationships among the North Gurage languages. He mentioned three shared features of

---

<sup>4</sup>East Gurage and Harari can only form relative clauses from PFV verbs; therefore, the REL IPFV is followed by the AUX *allə* which conjugates like a PFV and can take the relative clause marker.

these languages. (1) special suffixes that are added to the past tense copula - *u*. In North Gurage, this copula suffix is added to the verb, and for example, *banə + u* becomes *banə* for the 3.SG, but the copula remains the same, *banə* in Central West Gurage. (2) North Gurage languages have the Semitic root *hlf* ‘go’ while the Central West Gurage maintains *hwr* as in <sup>5</sup> *wərə* ‘to go’. (3) North Gurage languages have *-mu/-m* as plural markers.

Hetzron classified the tt-group into Muher and Western Gurage. Western Gurage is further classified into Mesqan (2TG) and 3TG (since they exhibit three tense markers). West Gurage languages are characterized by two features. (1) Reduced plural marker, *-(a)mu*, in Central Gurage reduced to *-o*. (2) They use imperfect as present only and developed two future tenses: definite and indefinite. Hetzron classified 3TG into Central West Gurage (CWG) and Peripheral West Gurage (PWG). Central West Gurage consists of Ezha, Chaha, Gumer and Gura while Peripheral West Gurage consists of Geyeto, Inor, Endegagn and Mesmes. (1) Peripheral West Gurage has ‘long nasality’ or ‘nasality spread’, nasal phoneme nasalizes all adjacent nasalizable phonemes: *arəmədə*, /r/ is nasal. (2) Loss of laryngeal sound in Central West Gurage and partial survivals in Peripheral West Gurage. (3) In Peripheral West Gurage, all laryngeal sounds except plain /h/ are lost with ruminant of long vowel, *\*dak’ə = daak’ə* ‘he laughed’ in Geyeto, and *daaʔa* in the rest. Finally (4), Peripheral West Gurage displays a survival of the Main Verb Marker in the past tense of copula, based on the stem *baanə* as in *baanə-da* where *-d* is a variant of the Main Verb Marker *-t*.

Though this classification is the most detailed and the comprehensive one as far as the classification of these languages are concerned, there have been critics about the classification, for instance, the parameters used are not uniform. Though Hetzron argued that morphological innovations are the best ‘isoglosses’, he used several other parameters including phonological and syntactic features (see Goldenberg, 1977).

Demeke (2001) re-examined this classification proposal and came up with an al-

---

<sup>5</sup>In this example, the first consonantal root is lost through time, and they are reduced from three to two.

ternative classification, with modification on the position of some of the languages. Demeke did not provide his own evidence for many of the arguments, rather he relied on the classification proposals of others (e.g., Leslau, 1969; Rose, 1996). Many of the disagreements between Hetzron (1972) and Demeke (2001) are on the classification of Gurage varieties as can be seen from Figure 3.2. They do not differ on the classification of South Ethiosemitic languages into Transverse South and Outer South. Both of them also classified Amharic and Argoba under Central South. However, there is a difference between them on the classification of Eastern South. Unlike Hetzron (1972), Demeke (2001) included Wolane under this group and presented as a sister language of Silt'e. Hetzron (1972) did not include Wolane in his classification probably because he was convinced that it is a dialect of Silt'e. It is also important to note that Hetzron (1972) considered Zay as a sister language to Silt'e (classified under East Gurage). Demeke (2001), however, considered it as a separate language (as a sister language to East Gurage). Nevertheless, Demeke (2001) did not provide convincing evidence for classifying Zay as a sister language of East Gurage. The classification of Zay under Eastern Gurage by Hetzron (1972), however, seems to be partly motivated by previous studies. Other scholars, for example Bender (1971) classified Gurage languages into three: East Gurage which consists of Silt'e, Wolane, and Zay; West Gurage which includes Chaha, Gyeto, Ezha, Inor, and Mesqan, and North Gurage which comprises only Kistane. However, neither of them provided adequate evidence to justify the position of Zay in their classifications. Demeke (2001) indicated that branching of Zay directly from the Eastern South branch even does not necessarily mean that Zay has closer relation with Harari than with the East Gurage varieties such as Wolane and Silt'e. He, nevertheless, suggested that East Gurage languages (Silt'e and Wolane) are very homogeneous as compared to both Zay and Harari.

Significant differences between the two works can be seen in the classification of the Outer South Ethiosemitic. Demeke (2001) rejected the -n and -tt classification arguing that these markers are not the features of all the languages that Hetzron (1972) mentioned. Instead, he created a group based on the suggestion of Leslau (1969). He took the initial letter of some of the languages in the group (Gogo,

Mesqan and Soddo (Kistane)) and created GMS-group, and he divided the Outer South Ethiosemitic into GMS-group and Western Gurage. However, he did not provide linguistic evidence for the formation of GMS-group. GMS-group consists of almost all the languages that are included in n-group in Hetzron's classification except Mesqan. In both cases, the group consists of Gafat and North Gurage. However, as Figure 3.2 depicts, the two scholars differ on the position of Mesqan. Hetzron (1972) classified Mesqan under West Gurage while Demeke (2001) classified it under North Gurage.

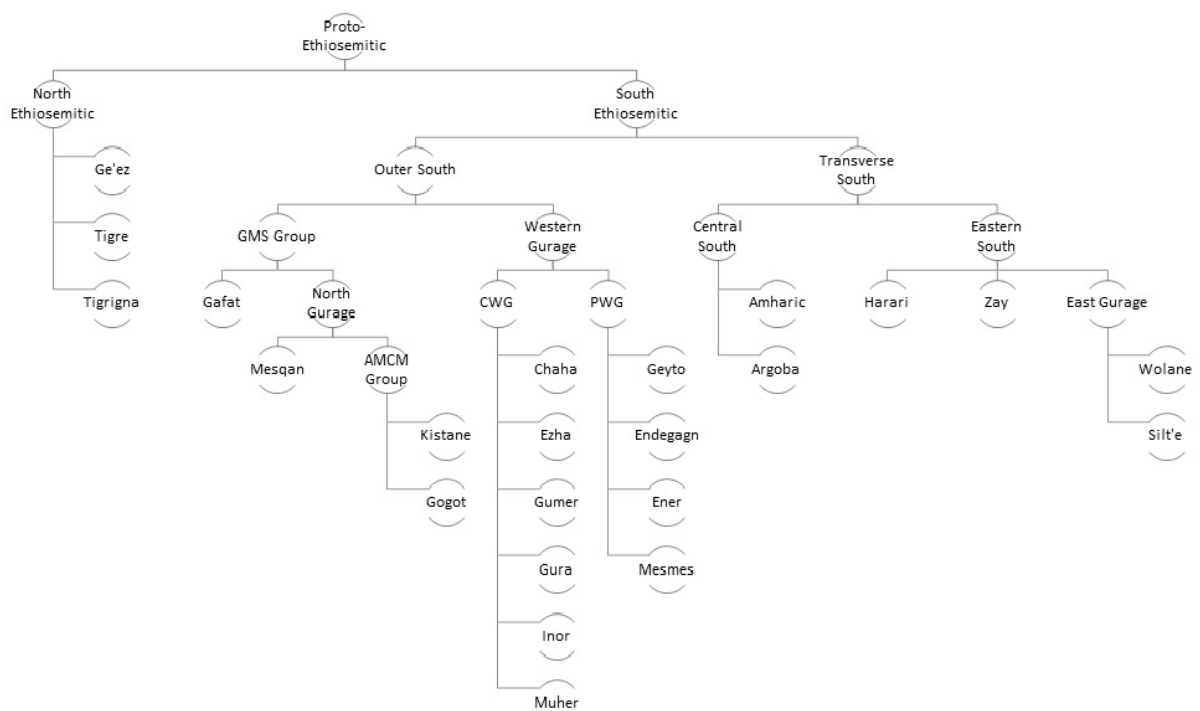


Figure 3.2: Classification of Ethiosemitic languages according to Demeke (2001)

With regard to the inclusion of Mesqan under North Gurage, Demeke (2001) provided four reasons. According to his argument, like other North Gurage languages (Kistane and Gogot), Mesqan has (1) only two tenses (past and non-past); (2) it has overtly marked simple tenses, but this is not true for Western Gurage languages because they either have a tense marker for simple past or for only future tense. (3) It has only present perfect construction with the perfective form of the verb and the suffix *-m*. (4) It has non-concatenative verbal pattern similar to that of Kistane and Gogot. However, Demeke (2001) did not consider Mesqan as a sister

language of Kistane and Gogot since it has slightly different features from the two. He indicated that Kistane and Gogot are more related compared to Mesqan since (1) Type B roots in Kistane and Gogot are known for gemination of penultimate consonants throughout the verbal paradigm and the occurrence of the front vowel after the first radical in perfective and in imperfective aspectual stems which is not the case in Mesqan. (2) In Kistane and Gogot, there are main verb markers which encode polarity but not in Mesqan. Hence, using this distinct feature he classified Kistane and Gogot under one group (AMCM) mainly due to affirmative main clause marker they contain.

As Figure 3.1 shows, Hetzron (1972) derived Muher and Western Gurage from the *tt*-group. However, Demeke (2001) dissolved the *tt*-group arguing that the *-tt* feature is not a representative of all the languages in the group, and he directly derived Central West Gurage and Peripheral West Gurage from Western Gurage. Clearly this change primarily affects the position of Muher. Demeke (2001) classified Muher under the Central West Gurage citing Leslau (1992) and Leslau (1969). He also argued that like other Central West Gurage languages, Muher has simple past with the morpheme *-m* attached to the perfective stem and the main verb markers either in the present tense or in non-past tense. He also indicated that Muher is similar to other Central West Gurage languages in terms of its penultimate consonants in perfective paradigm. Hence, Muher moved to Central West Gurage along with Ezha, Chaha and Gura. However, according to Demeke (2001), unlike other Central West Gurage varieties, Muher has a visible morphological marker for the present tense.

Demeke (2001) also rejected the 3TG classification of Hetzron (1972), arguing that this tense type does not exist in all the languages. Hence, he divided Western Gurage into Central West and Peripheral West Gurage. This division is the same both for Hetzron (1972) and Demeke (2001). However, the two differ on the position of Inor. Unlike Hetzron (1972), Demeke (2001) classified Inor under Central West. Quoting Leslau (1996), he argued that Inor is close to Central West Gurage languages particularly Chaha. He further supported his argument by citing the data from his informant. According to his <sup>6</sup>informant, Inor is similar to Chaha.

---

<sup>6</sup>Demeke's (2001) analysis was predominately based on the previous works and informants in

### 3.3 Summary: Issues in the Classifications of Ethiosemitic Languages

From the two classification proposals and the works of others presented above, it is obvious that the classification of Ethiosemitic languages is not a finished business. As indicated in 3.2, though the classification of Ethiosemitic languages into North and South is generally accepted by many scholars, there are yet linguists who suggest (e.g. Fleming, 1968; Gragg, 19997; Voigt, 2009) further investigation. Besides, as presented above, Hetzron (1972) and Demeke (2001) have disagreed on the classification of some of the languages. Zay is one of them. Demeke (2001) classifies it under Eastern South Transverse while Hetzron (1972) classifies it under East Gurage. The two classifications are very different since Hetzron considers Zay as a Gurage language, but Demeke does not. Other scholars such as Meyer (2005a), Meyer (2006), Gutt (1980) and Hudson (2000) also recognize that Zay is East Gurage language, mainly probably based on the works Hetzron (1972) and other studies. This disagreement indicates that further investigation should be done on Zay and other East Gurage languages to determine the exact position of Zay.

As presented in 3.2, there has been a disagreement among scholars regarding the genealogical relationship among the Gurage language varieties. According to Fleming (1968) and Faber (1997), Gurage languages which constitute a large number of the South Ethiosemitic languages, do not seem to have a common genealogical relationship. For instance, East Gurage is closer to Harari than to the rest of the Gurage languages. Furthermore, Kistane and Gogot are closer to Gafat than to other varieties. The relationship between Zay and Harari is also not clear. Demeke (2001) considered them as sister languages, but Hetzron (1972) did not. Whether Wolane is a language or a dialect of Silt'e is another area that needs further scrutiny. Meyer (2006) recognizes that the two varieties are mutually intelligible based on descriptive linguistic grounds, but he also prefers to consider the two varieties as separate languages based socio-linguistic grounds; the speakers of the two varieties

---

Addis Ababa. It appears to me that he did not conduct fieldwork for many of the languages.

believe that there are cultural differences between them. Demeke (2001) treated Wolane as a separate language while Hetzron (1972) considered it as a variant of Silt'e. Indeed Demeke also agreed that Wolane is very similar to Silt'e. Other scholars such Gutt (1980) and Menuta (2015) also agree that Wolane is very similar to Silt'e.

As presented in 3.2, the classification of the Outer South Ethiosemitic languages is the most unsettled one. The classification into n-group and tt-group by Hetzron (1972) was later reshuffled by Demeke (2001) since these tense markers are not shared among the languages that are included in each of the categories. The alternative classification into GMS-group and Western Gurage which is proposed by Demeke (2001) is supported by some evidence but not detailed enough. The two scholars also do not agree on the position of Mesqan. Hetzron (1972) classified it under West Gurage while Demeke (2001) classified it under North Gurage. Though Demeke (2001) provided some evidence (see §3.2) for doing so, it is obvious that the detailed description of shared innovations of all the languages in both West and North Gurage is required to determine the exact position of Mesqan. The relationship between Gafat, Kistane and Gogot also need further investigation. Gafat has a remote relationship with the two languages in Demeke's classification than it has in the classification of Hetzron (1972).

Moreover, Muher does not seem to have a settled position in the classification of the South Ethiosemitic languages. Hetzron (1972) classified it under tt-group. However, Demeke (2001) moved it to the Central West Gurage. It is important to note that Central West Gurage is the mother language of Muher in Demeke's classification, but it is a sister language in Hetzron's classification. Since Demeke was also not really certain about the exact position of Muher, determining the position of Muher in South Ethiosemitic languages is another area that needs further investigation. Both scholars also differ on the position of Inor. Demeke (2001) categorized it under Central West Gurage while Hetzron (1972) classified it under Peripheral West Gurage. Neither of them provided detailed and convincing evidence for their classification.

As indicated in the introduction, the classification proposals presented through-

out this chapter were mainly based on morphological parameters. Due attention was given to these works since they are the major resources available with regard to the classification of Ethiosemitic languages, and they are indeed vital to illustrate the core discussion points with regard to the classification of Ethiosemitic languages. There are some classification attempts based on a lexical parameter. These studies will be discussed in the next chapter.



---

## Chapter 4

# Measuring Linguistic Distance: the Approaches

This unit presents an overview of the main theories of language variation (§4.1), the perspectives that the previous studies have taken in the attempt to determine the distance among closely related languages (§4.2), the relationships among various perspectives of measuring linguistic distance (§4.3), and major factors that influence language change and variation (§4.4). The chapter aims to highlight major theoretical and methodological assumptions that are useful for the argumentation about the classification of Ethiosemitic languages.

### 4.1 Theories of Language Variation

Before 19<sup>th</sup> c, there was no a concrete explanation about the sources of language diversity (Geisler & List, 2013, p.113). A discussion regarding the origin of languages was limited to the Biblical myth of the Tower of Babel. Nonetheless, starting from early 19<sup>th</sup>c, there have been two conflicting models: Tree model and Wave model (Hamed, 2005, p.1015; Heggarty, 2010, p.301). The tree model (Schleicher, 1853) was inspired by the Darwinian theory of evolution of species. It was built mainly based on sound changes. According to Geisler & List (2013, p.114), the Tree model supposes that sound change is a universal, gradual and law-like (exceptionless) process. Following this model, as has been the case in the differentiation of genome

in biology, in linguistics, binary trees which indicate the splitting of the ancestral into child languages have been used to illustrate historical changes of the languages (Geisler & List, 2013, p.111; François, 2015, p.163).

However, according to Geisler & List (2013, p.113) and François (2015, p.165), the tree model was continuously disputed in linguistics. The Tree model fails to deal with hidden borrowings in languages which are not genetically connected. Moreover, in the Tree model, languages are treated as discrete objects located in space and time. It assumes a complete separation of languages and neglects the horizontal relationship among the languages. In other words, the Tree model focuses only on the vertical relationship. Given these limitations, Schmidt (1872) argued for non-hierarchical diffusion of linguistic innovations which can be because of borrowing or creolization. He proposed an alternative model, the Wave model, which accounts for the areal diffusion. According to the Wave model, changes in linguistic features spread as waves in concentric circle over neighboring speech communities. The Wave model adequately addresses the spatial expansion of languages which was neglected in the Tree model, but it addresses the contact-related issues by neglecting the vertical relationship among languages (François, 2015, 167-168; Zobl, 1984, p.160-161; Bailey, 1973, p.67).

It seems that the two models can be best utilized in different contexts. According to François (2015, p.170), whereas the Tree model is adequate for most studied language families such as Indo-European, there are innovations which cannot be explained by the Tree model. In most cases, the two models are kept apart; the Wave model is used to explain contact-induced language changes whereas the Tree model is assumed in the explanation of historical relationships among languages. Hence, historical linguists often try to circumvent contact-related situations by selecting linguistic features that are immune to borrowing. However, in most cases, such an attempt to control borrowing is not fruitful. Usually different language varieties are mixed and lead to a homogenized context, contrary to tree-like development. According to Geisler & List (2013, p.112) and Zobl (1984, p.178), the most plausible alternative could be to combine both models in a network approach where both vertical and horizontal relations are displayed. This approach preserves the advan-

tage of the Tree model's dichotomous logic with clear-cut categorization, but further allows fine-grained mapping of language contact (Geisler & List, 2013, p.122). The debate is still ongoing. A recent rather radical approach is the claim that the Tree model is just an instance of the Wave model (see Agee, 2018; François, 2015; Kalyan et al., 2018). These studies argue that the Wave model, combined with the comparative approach of historical linguistics, can best explain the historical changes among related languages. A discipline associated with this approach, *Historical Glottometry*, has becoming very popular. There are other models which are in one way or another associated with the Wave model. The Gravity model (Labov, 2001; Trudgill, 1974) is one of them. The Gravity model supposes that linguistic diffusion is determined by the combination of social gravity and geographical distance (see §4.4.2.1). It further argues that the diffusion of linguistic features does not necessary involve a continuous expansion of linguistic features. Rather, there is always unaffected intermediary locations.

Tree and Wave models explain just the general principles underlying the formation of new languages. There are other fine-grained hypotheses associated with the formation of new languages or dialects. For example, Trudgill (2004, p .88-120), in his *New Dialect Formation* proposal, identified three major stages of new language formation. According to this proposal, the first stage is rudimentary leveling and interdialect development, stage two is variability and apparent leveling, and stage three is the survival of the majority forms. In the same manner, in his *Dynamic Model*, Schneider (2003) identified five stages in the formation of English dialects across former English colonies. These stages are determined based on both linguistic and sociolinguistic (identity reconstruction) parameters. The stages include *foundation* (regular uses or the introduction of a new language in a new context and subsequent contact), *exonormative stabilization* (political stabilization with the dominance of the colonizers, and sense of different linguistic identity), *nativization* (the full confirmation of identity), *endonormative stabilization* (political independence and linguistic standardization) and *differentiation* (the emergence of a new variety). The fundamental difference between Dialect Formation model and the Dynamic model is that the later emphasizes the importance of non-linguistic factors in

the process of new dialect formation. In the same manner, Britain (2017) discusses various non-linguistic factors that contribute to language change and to new dialect formation. The major factors are population size, duration of contact, degree of similarity among the languages, degree of social integration, landscape, attitude of the speakers and many others. Non-linguistic factors that contribute to language change will also be discussed in 4.4.2.

## 4.2 Measuring Linguistic Distance

From the discussion in 4.1, it is clear that the distance among related languages is the outcome of the dynamics within the language themselves and the influences of several extra-linguistic factors. In the sections that follow some of the methods that have been employed to measure the linguistic distance are discussed. Discussing all the methods of measuring linguistic distance is beyond the scope of the present study. Hence, among various methods of measuring structural distance, only the most common ones, i.e., traditional (§4.2.1), computation (§4.2.2), functional (§4.2.3) and perceptual approaches (§4.2.4) are the focal points of this section.

### 4.2.1 Structural Distance - Traditional Approach

The structural distance is determined based on the degree of phonological, morphological, lexical or syntactic difference among related languages. There are different traditional methods of objectively measuring linguistic distance. Some of them are discussed as follow.

#### 4.2.1.1 The Isogloss Method

An isogloss is a line on a map which divides areas where dialects differ (Kretzschmar, 1992, p.227). According to Heeringa (2004, p.10), the use of isogloss for the classification of language areas was derived from an *isotherm*, a line on a map that connects areas that have similar temperature in *meteorology*. During dialect classification, isoglosses of different phenomena are drawn on a map, and the coinciding isoglosses

are taken as a border line between the adjacent dialects. The isogloss method determines language boundaries usually based on the similarities or differences between linguistic features; the parameters are in most cases linguistic parameters. However, according to Kessler (1995, p.60) and Heeringa (2004, p.10), using isoglosses has three weaknesses. First, isoglosses of different features do not always coincide. They can be parallel or even can sometimes cross each other which leads to the creation of a vague and contradictory division of dialects. Even when they coincide, the isogloss method imposes the requirement of reduction to one dimension; it imposes the characterization of a certain geographical area by a unique dialect. However, linguistic features usually cannot be separated by a sharp geographical boundary.

Moreover, often the variants do not lineup on two sides of a line, but they are intermixed to a certain extent. Usually adequate information may not be available for certain sites or dialects. It can also be the case that certain linguistic variables are irrelevant for some dialects. Besides, the isogloss method conceptually presupposes drawing a clear dialect boundary where in reality there is a continuum with gradual dialect changes (Heeringa, 2004, p.10; Rabanus, 2017, p.362). More importantly, according to Goosens (1977, 1977, p.56) and Heeringa (2004, p.10), isoglosses cannot be drawn without making a subjective judgment.

#### 4.2.1.2 The Structural Geographic Method

The term *Structural Geographic Method* was used in Heeringa (2004). It is used to classify language areas based on the structure of the languages. The structure may refer to any language feature: phonetic, lexical, morphological or syntactic features. This means that dialects which share the same phonological, morphological, syntactic or lexical features can form a dialect area (Heeringa, 2004, p.9). For instance, Moulton (1960) classified the dialects of the northern Switzerland based on short vowel system. To add more, Amharic is often divided into Shewa, Gojjam, Wello and Gonder dialects based on either the lexical, phonetic or grammatical differences (see Marcos, 1970; Tadesse, 2018; Zelealem, 2004). Though the Structural Geographic Method is the most commonly used method of dialect classification, it has its own weaknesses, i.e., it focuses only on one dimension of language differences although

languages in reality differ in multiple dimensions. For example, the classification of dialects just based on phonetic differences may not be adequate to classify related languages comprehensively since different dialects may have similar phonetic features, but differ in many other dimensions (Heeringa, 2004, p.10).

## **4.2.2 Structural Distance - Computational Methods**

Different computational methods of measuring linguistic distance have been developed in the last few decades (see Heeringa, 2004, p.123-200). Many of these methods classify languages mainly based on their phonetic or lexical differences. Some of these methods are presented below. A special attention is given to the Levenshtein distance and lexicostatistics for two reasons. First, they are the most extensively employed methods in dialectometry. Second, they are core parts of the methods used in the present study.

### **4.2.2.1 Counting Differences and Similarities**

The counting differences and similarity method is one of the oldest computational approaches. As the name indicates, it involves counting the differences and the similarities between related or adjacent dialects. As a general principle, a questionnaire containing various deliberately selected language features, for example, phoneme inventories is provided to dialect speakers. Then, the number of items upon which the speakers of two or more dialects disagree is presented in percentage, and this percentage is considered as an index score indicating the distance between the dialects (Heeringa, 2004, p.12-16). For example, Séguy & Ravier (1973) counted the number of items on which the speakers of neighboring dialect disagreed. The number of disagreements between the two neighbors was expressed in terms of percentage, and the percentage was considered as an index score which shows the linguistic distance between two or more dialects.

#### 4.2.2.2 Corpus Frequency Method

According to Heeringa (2004, p.16-20), Corpus Frequency method was introduced by the Hoppenbrouwer brothers (1988). There are two different methods within Corpus Frequency method: Letter/Phone Frequency Method (L/PFM) and Feature Frequency Method (FFM). In the Letter Frequency Method, unigram frequencies of letters are computed based on a sample corpus. Frequencies are computed for different letters in each of the related languages, and the differences between the corresponding letter frequencies is taken as the distance between two corresponding dialects or languages. The frequency is normalized (converted from absolute frequency to relative frequency) since different corpora may not have the same size. The main limitation of Letter Frequency Method is that usually there is no one-to-one correspondence between the phoneme and grapheme.

Because of this limitation, the shift was made by the two brothers from computing the letter frequency to computing the phone frequency. The procedure of the frequency computation is similar to that of Letter Frequency except that in this case IPA transcription is employed instead of letters. According to Heeringa (2004, p.16), the two brothers employed the IPA transcription of 'The North Wind and the Sun' to compute the distance. However, the phone frequency method has also its own limitation. The major shortcoming of the method is that it does not precisely illustrate the degree of difference between phones. Rather, it entirely depends on numerical values which are sometimes misleading (Heeringa, 2004, p.16-20). To elaborate this point further, I repeat examples used by (Heeringa, 2004, p.19). Let us assume that there are three dialects: Dialect A, Dialect B and Dialect C, and these dialects are compared to each other in terms of the frequencies of three vowels: [e], [i] and [u] and let us suppose that the relative phone frequency computation yielded the following results.

Table 4.1: Phone Frequency Method

<b>Dialects</b>	<b>[e]</b>	<b>[i]</b>	<b>[u]</b>
Dialect A	100%	0%	0%
Dialect B	0%	100%	0%
Dialect C	0%	0%	100%

If we compute the distance between pairs of the dialects indicated in Table 4.1, we find similar distance among all the languages. However, the actual distance among the three dialects may not be precisely the same especially if the features (e.g., manner and place of articulation) of the vowels are strictly considered. For example, front vowels such as [e] and [i] are more similar to each other as compared to front and back vowels such as [e] and [u] respectively.

Because of this limitation of the phone frequency method, the two brothers introduced the Feature Frequency Method (FFM) (Heeringa, 2004, p.19-20). Feature Frequency Method considers the frequency of the features of the vowels and the consonants rather than focusing just on the frequency of the phones. In other words, it takes into account specific features of sounds based on the place and manner of articulation, i.e., [+voice, -voice], [+aspirated, -aspirated], [+nasal, -nasal], [open, close], [back, front]...etc. According to Heeringa (2004, p.19), the weakness of all frequency methods is that they do not attach any significance to the word. In other words, the difference in the features of frequency of the vowels or the consonants between two dialects say nothing about difference between the two dialects at the word levels. Put simply, this method neglects the lexical differences that occur due to the dialectal variation. Due to this limitation, the Frequency Per Word Method was introduced.

#### 4.2.2.3 Frequency Per Word Method

Frequency Per Word Method is similar to the Feature Frequency Method. However, it compares features of phones within pairs of words, not within the whole corpus. This means that the comparison is narrowed from corpus level to word level; the method considers words as separate entities (Nerbonne & Heeringa, 2001, p.73). The



comparison between two or more words can be made either by computing the phones or the feature frequencies which are usually expressed in terms of percentages. In order to use this method, words in selected dialects/languages are first transcribed. The distances between them are the average of the phone difference/similarities observed in each of the transcribed pairs of words. In other words, the dialect distance is computed by dividing the sum of the distances between each pair of words by the total number of the pairs of words. On the bases of this computation, the distance matrix of the dialects can be obtained. Then, cluster analysis is performed on the matrix. Since Frequency Per Word Method does not take into account the sequence of phones in pairs of words, it assumes, for example, that English ‘tool’ [tu:l] and ‘loot’ [lu:t] have zero distance. This limitation led to the emergence of Levenshtein distance (Heeringa, 2004, p.20; Nerbonne & Heeringa, 2001, p.73).

#### 4.2.2.4 Levenshtein Distance

The Levenshtein distance was first introduced by Levenshtein (1966). It is a string edit distance measure which calculates the distance between the pronunciations of corresponding words. In another words, it is “a numerical value of the cost of the least expensive set of *insertions*, *deletions* or *substitutions* that would be needed to transform one string into another” (Heeringa, 2004, p.23; Kessler, 1995, p.63). Unlike the Frequency Per-Word Method, Levenshtein distance takes into account the order of phones in a word.

##### 4.2.2.4.1 The Basics of Levenshtein Distance

Levenshtein distance computes a minimal cost required to transform one pronunciation into another pronunciation by means of insertions, deletions or substitutions (Beijering et al., 2008, p.13; Gooskens, 2007, p.455; Heeringa, 2004, p.23; Kessler, 1995, p.63). According to Heeringa (2004, p.5), employing Levenshtein algorithm to measure pronunciation distance was introduced by Kessler (1995). He applied the algorithm to measure the distances among pairs of Gaelic dialects. The advantage of the Levenshtein distance is that it takes the order of phones into account. Another

major advantage of Levenshten distance is that it considers the length differences between strings or words to be compared. For example, an Amharic word ‘power’ can be pronounced as [textithak’im] around Gojjan, but as [ak’im] around Addis Ababa; hence, four sounds vs. five sounds. Levenshtein algorithm is capable of handling this kind of length of phone differences between pairs of words. Furthermore, regardless of many different operations that map the production of any of two words, Levenshtein distance always looks for the ‘cheaper’ one, in terms of the effort (cost) which is necessary in order to transform one sequence of phones into another one.

The notion of minimal cost and string operation are central to the concept of Levenshtein distance. Hence, the two concepts are elaborated below with illustrations. The following examples represent the pronunciations of an Amharic word, ‘gathering’. The word is pronounced in slightly different manner in two locations in Ethiopia: Wello and Addis Ababa. The word is taken from Marcos (1973, p.128). The phonetic distance between different pronunciation of this word can be computed as follow.

Table 4.2: Examples of Levenshtein distance

s	i	b	s	a	b	o	<b>Wello Dialect</b>
s	i	b	s	ə	b	a	<b>Addis Dialect</b>
			1		1		<b>cost = 2</b>

The operation that transformed the pronunciation of [sibsabo] to [sibsəba] can be presented as follow.

Table 4.3: Examples of Levenshtein distance

s	i	b	s	a	b	o	<b>substitute a by ə 1</b>
s	i	b	s	ə	b	o	<b>substitute o by a 1</b>
s	i	b	s	ə	b	a	
<b>cost</b>							<b>2</b>

The three common operations of the Levenshtein distance are deletion, insertion and substitution. The first two are often called *indels*-the blending of *insertion* and *deletion* (Heeringa, 2004, p.123). The Levenshtein distance can also be denoted

as the least cost that maps the pronunciation one word,  $w_1$  to the pronunciation of another word,  $w_2$  (see Heeringa, 2004, p.123). In other words, the Levenshtein distance can be expressed as a function of the distance between the pronunciations of the two words,  $d(w_1, w_2)$ . As a general rule, for substitution, the weight is equal to the distance between the corresponding segments. For indels, the weight is equal to the distance between the inserted or deleted phones and ‘silence’. Since words or strings have different lengths, the distance between pairs of words are sometimes normalized. The distance between two words is said to be normalized when the total cost is divided by the longest alignment. In Table 4.2 and 4.3, both pronunciations have equal length, seven segments. The Levenshtein distance is 2. Then, two divided by seven gives the normalized Levenshtein distance which is 0.28, and 28% in terms of percentage (Heeringa, 2004, p.130). The distance between languages or dialects is assumed to be the mean of the distance between all pairs of words/cognates in a sample speech/corpus.

To compute the Levenshtein distance, various software packages can be employed. Several studies which were conducted on Germanic languages employed RuG/L04, ‘linux-flavored’ package, developed by a group of researchers at University of Groningen (see Nerbonne et al., 2011). Since, RuG/L04 requires command-line commands which is very complex for researchers that do not have a good computer experience, recently a group researchers at the University of Groningen has developed another user-friendly graphical user interface which is called GabMap, hence, it can be used by people of all kinds of backgrounds. The features and the procedures of using GabMap can be referred from Nerbonne et al. (2011). An alternative tool is DiaTech, widely used in the study of Romance languages (see Aurrekoetxea et al., 2013).

Levenshtein distance can be applied to three types of data sets: transcribed words, sound features and acoustic signals (see Heeringa, 2004, p.121). They are summarized as follow. To begin with the transcription of words, the Levenshtein algorithm can be used to compare the transcribed pairs of words or cognates, and in fact this is the most common approach that has been used in many of the previous studies. This approach ignores supra-segmental features during phone comparisons.

The extent to which one pronunciation differ from another is determined by computing the difference between segments. Most often, during the comparison of the transcribed pairs of words, equal cost is assigned to each operation - 1 point, which means that a slight difference between very similar phones, for instance bilabials [p] and [b] and that of quite distinct phones, for example, velar [k] and back vowel [o] receive the same cost. Diacritics such as nasalization are often ignored (Kessler, 1995, p.63).

The weight assignment can also be flexible, where different costs can be assigned to different operation based the nature of the sounds to be compared, and partly based on the degree of precision desired in a study. Another essential concern with regard to measuring linguistic distance based on the transcribed words is whether all the words in a sample speech or only the phonetic variants of cognates should be considered (Heeringa, 2004, p.34). Kessler (1995) applied both approaches on Irish dialects. In the terminology of Kessler (1995), the first one is called *all-word level approach*, and the second one is *same-word approach*. Many of the previous studies computed just the distance among cognates (e.g. Heeringa & Nerbonne, 1999; Gooskens, 2007; Van Bezooijen & Gooskens, 2007). Related to this, Kessler (1995, p.64), suggests that there is no need to compare words that are historically not related. In the same manner, according to Nerbonne & Heeringa (2010, p.555), Dutch dialect source which was analyzed in two ways: once restricted to cognate words, and once comparing all semantically similar words, including unrelated words showed almost perfect correlation ( $r = 0.98$ ). When just cognates are used for the computation of linguistic distance, the percentage of cognates that should be used need to be determined. For instance, in Heeringa & Nerbonne (2000), cognates that are shared by 60% of the languages were used for phonetic comparisons. The details of the computation of Levenshtein distance based on the transcribed pairs of words can also be inferred from Heeringa (2004, p.121-135).

The second approach is using Feature Representation. In this case, Levensthtein distance is employed to compare the features of sounds, not the phones themselves. The characteristics of vowels and consonants are expressed in terms of some defined features based on their manner and place of articulation. In other words,

each phone is represented by collections of features. Then, a discrete ordinal value that ranges between 0 - 1 is assigned to each of the corresponding feature. The assignment of the value is usually arbitrary. Nonetheless, the weight assigned to each feature can also vary depending on the nature of the features to be compared (Heeringa & Nerbonne, 2004, p.20-62; Kessler, 1995, p.63-64). For instance, Kessler (1995) used twelve features: nasality, stricture, laterality, articulator, glottis, place, palatalization, rounding, length, strength, and syllabicity to compute the distance between Gaelic dialects. The number and type of features to be compared is largely dependent on the nature of the phonology of the particular languages. The distance between any two phones is the difference between the feature values<sup>7</sup>, averaged across all the twelve features. For instance, the distance between [k] and [b] in a given word can be computed as follow.

Table 4.4: Feature Representation

<b>Features</b>	<b>[k]</b>	<b>[b]</b>	<b>Distance</b>
Glottal	0	0	0
Post Alveolar	0	0	0
Velar	.3	0	.3
Pre-velar	0	0	0
Palatal	0	0	0
Pre-palatal	0	0	0
Alveolar	0	0	0
Dental	0	0	0
Labial	0	1	1
<b>Mean Distance</b>			<b>1.3/9 = 0.14</b>

The obtained matrix is considered as feature string comparison. In Kessler (1995), the feature string comparison was made only between pairs of citations that used the same word/cognates. The comparison of the features can be made using feature histograms, and Pearson correlation. The simplest way of computing

<sup>7</sup>The cost assignment used by Kessler (1995)) was as follow: Glottal = 0, uvular = 0.1, postvelar = 0.2, velar = 0.3, prevelar = 0.4, palatal = 0.5, alveolar = 0.7, dental = 0.8, and labial = 1. This cost assignment is arbitrary based on the individual's preference.

the distance is calculating the absolute value of the difference between each of the corresponding feature values. The feature representation approach is particularly important since the comparison of pairs of transcribed words does not take into account the difference between very similar segments such as aspirated [p<sup>h</sup>] and unaspirated [p] (Heeringa, 2004, p.34). It is very important to mention here that this is a very shallow and extremely condensed presentation of Feature Representation approach. For detailed explanation about this approach, interested readers are referred to Heeringa & Nerbonne (2000) and Heeringa (2004).

The third approach is using <sup>8</sup>Acoustic Segmental distance measure. In this case, the Levenstein distance is applied to measure the distance between two varieties based on the acoustic properties of the sample speeches/sounds: energy concentration in the spectrogram. Spectrograms are made for the sample speech sounds. The spectrogram is a graph containing frequency (on the vertical axis), time (on the horizontal axis) with darkness of the graph at any point representing the intensity of the sound.

Wieling et al. (2012) argues that cost assignment approach which employs weighting schemes based on the inverse frequency with which segments are aligned is superior to the binary cost assignment. However, according to Heeringa (2004, p.5), any of the three approaches can be used based on one's own interest and research objectives. For instance, if the goal is to approximate how difference among the languages are perceived by dialect speakers, the use of binary cost outperforms the gradual cost. This implies that segmental difference is more important than the degree to which they differ. Furthermore, according to Heeringa et al. (2006, p.60), the choice between relative and absolute distance is dependent on the research objectives. Absolute Levenshtein distance approximates dialect difference as perceived by the native speakers better than the relative Levenshtein distance. Furthermore, according to Gooskens et al. (2008, p.63), mutual intelligibility among languages correlates better with relative distance that with the absolute distance.

---

<sup>8</sup>This method does not try to automatically recognize the sounds contained in the signal. It just compares the spectrogram structure.

#### 4.2.2.4.2 Levenshtein Distance in Previous Studies

In 4.2.2.4.1, the basic concepts of the Levenshtein distance were presented. In this section, some of the studies that employed the Levenshtein distance and the procedures employed in the studies are summarized. In one way or another, these studies employed one or more than one of the three Levenshtein distance approaches presented above. A computational approach of language classification in general is a new phenomenon in the context of Ethiopian languages. As far as my reading is concerned, Negesse (2015) is the only study that applied such an approach to the classification of Oromo dialects. Hence, almost all the studies reviewed below were conducted on non-Ethiosemitic languages.

Several studies previously employed Levenshtein distance to compare the transcriptions of pairs of words. However, the nature of the transcriptions used and the degree of restrictions imposed on the transcriptions varies from study to study. For instance, Van Bezooijen & Gooskens (2007) employed the simplest variant of Levenshtein distance to determine the distance between written Dutch, Frisian and Afrikaans. The distance measure was between written forms (letters) rather than between the transcriptions of words. The comparison was between cognates taken from a fable ‘the North Wind and the Sun’. Equal weight was assigned for each operation, 1 point. The distance between each pairs of cognates was normalized. Gooskens (2007) used Levenshtein algorithm to determine the relative distance between three Scandinavian languages: Danish, Norwegian and Swedish. Cognate words were taken from the recording of a story of ‘A kangaroo running on the streets of Copenhagen’ and transcribed. Unlike the equal cost assignment used in Van Bezooijen & Gooskens (2007), in this study different cost unit was assigned for the three operations based on the relative distance between the phones that are inserted, deleted or substituted: insertion and deletion 1 point, identical symbol 0 point, substitution of vowel by vowel or consonant by consonant 0.5 point, substitution of vowel by consonant or consonant by vowel 1 point and the presence or the absence of diacritics 0.25 point.

Heeringa & Nerbonne (2000) used the Levenshtien distance to measure the di-

achronic and synchronic distances between Frisian and Dutch dialects. They employed a parable ‘The prodigal son’, which contains about 186 words. The objective of the study was to determine the distance among the varieties based on the old and the new versions of the parable. To employ the Levenshtein distance, they selected 60 words which are shared both in the old and new parables (the same parable with slight difference due to changes in time). The cost assignment for the operation was somehow different. In this case, higher cost was assigned for substitution, 2 units and the remaining two costs, insertion and deletion each received 1 unit - the reason was not explained (see Heeringa & Nerbonne, 1999, p.90-92). Based on these distance matrices, the dialects were classified using cluster analysis and multidimensional scaling.

Beijering et al. (2008) employed the combination of acoustic segmental presentations and the phonetic transcription of cognates to determine the distance among 18 Scandinavian varieties. They compared cognates which were obtained from the fable ‘The North Wind and Sun’. The Levenshtein algorithm was adopted in such a way that vowels align with vowels and consonants with consonants. Equal cost (1 unit) was assigned for every operation. Moreover, spectrograms were created for the vowel and consonants. Based on the spectrograms, the distances between the vowels and pulmonic consonants were determined. They employed Barkfilter representation to measure the acoustic distance between the sounds. In general, there are several other studies that used Levenshtein distance. Nerbonne et al. (1999), Nerbonne & Hinrichs (2006), Serva & Petroni (2008), Schepens et al. (2012) and Silvestri & Tomezzoli (2005) are just some of them.

#### **4.2.2.5 Lexicostatistical Approach**

Lexicostatistics is a method of comparative/historical linguistics that involves comparing the percentage of lexical cognates between languages to determine their relationship. Lexical distance can be defined as the degree of similarity or difference among related languages on the bases of lexical comparison. Previous studies, largely relied on Swadesh word lists. Swadesh (1950) introduced a word list of 200 items which he assumed represent the core basic vocabulary of human language.



For the sake of convenience, he later reduced to 100 word lists (Swadesh, 1955). Since then, these lists of words have been used for various research interests; mainly for genealogical reconstruction of languages and estimation of the chronology of language diversification or *Glottochronology* (see Heggarty, 2010; Petroni & Serva, 2010; Swadesh, 2010).

Regardless of its magnificent contribution in the study of languages of the world, the Swadesh word list has entertained sustained criticisms. One of the problems is that some languages have been noted to lack equivalents for some of the items (see Alexei et al., 2010). Besides, the selection of the words was purely intuitive, and there is no objective criteria that make them the best candidates. More importantly, there has been a challenge of using the word lists for historical reconstruction of languages since there is no precise way of distinguishing borrowed words from those that are historically inherited (see April et al., 2010 for the detailed discussion). With all its limitations, lexicostatistics has remained one of the most commonly used approaches to measure the distance among related languages not only from diachronic perspectives but also from synchronic point of view.

#### 4.2.2.5.1 Basic Concepts of Lexical Distance

In a standard lexicostatistical approach, the lexical similarity of two or more languages is calculated by counting the proportion of shared cognates of some set of pre-selected words. Languages that have a greater lexical similarity are considered to be genetically more closely related than languages that have a lesser lexical similarity (Wang & Minett, 2005, p.127). The standard lexicostatistical approach also assumes a regular and gradual sound change. Hence, historical linguists usually apply lexicostatistics to determine a vertical relationship among languages using split trees.

Nevertheless, as noted above, there are theoretical problems related to the standard lexicostatistical approach. To begin with, language change is not always regular and gradual. Rapid and an unexpected language change, as a result of contact, for example borrowing, is possible. Second, from a practical point of view, there are

may concepts that do not have lexical representations in some of the languages. For instance, a Gurage word *koc'c'o* 'name of traditional food' does not have English equivalent. Contrary to this, a single concept can be represented by multiple lexical items (see Nerbonne & Kleiweg, 2003 for discussion on multiple responses). Often, there are also problems related to identifying shared cognates among two or more related languages since, in principle, this requires the full understanding of the history of each of the words. Given this, in previous studies, some scholars tirelessly worked to identify 'genuine' cognates using a very conservative approach (e.g. Bender, 1966; Fleming, 1968). Some others followed a very liberal approach in which cognates were identified based on a few shared sound inventories (e.g. Bender et al., 1976). Some were extremely liberal to the extent that they did not care about borrowed words (e.g. Hudson, 2013); hence cognates and borrowed words were not distinguished. Indeed, cognate identification can be extremely difficult especially when someone investigates the least studied languages that do not have dictionaries. In general, it appears to me that borrowing becomes an issue when someone aspire to trace the historical relationship among languages based on their lexical similarity. Scholars whose works focus on areal classifications of languages are less likely to be concerned about borrowing. Hence, there is a recent general tendency of drifting way from entirely relying on Swadesh word lists and using randomly selecting words. Some of the studies that followed this approach will be discussed in the next section.

#### 4.2.2.5.2 Lexicostatistical Studies on Non-Ethiosemitic Languages

As indicated above, many of the previous studies used the basic Swadesh word lists both to classify related languages and to estimate the historical time of separation of genetically related languages. Since the genealogical classification of related languages is not the focus of the present study, studies that employed lexicostatistics for areal language classification are emphasized here. Several recent studies compared lexicons that are drawn from randomly selected small corpora such as 'The North Wind and the Sun'. Based on this corpus, the percentage of the cog-

nates within two or more languages is computed to estimate the degree of similarity among closely related languages. It appears to me that historical linguists and dialectologists seem to differ on data sampling procedures. Historical linguists usually use carefully selected 'borrowing-immune' data for historical comparisons while dialectologists rely on randomly selected linguistic data, i.e., list of words. Besides, the procedure of cognate identification and what was really considered as 'cognate' differ from study to study to the extent that they are often controversial. For instance, in Tang & van Heuven (2007), a so-called Lexical Similarity Index (LCI) was used (adopted from Cheng (1997)) to measure the relative distance among the 15 Chinese dialects. According to their definition, the lexical similarity index is the association of coefficient *phi*, which indicates how much the words expressing a concept are shared among the varieties. For instance, if there are six words expressing a concept and if all these words are used in all varieties, the phi value is 1. However, if none of lexical items that is used in one variety is used in another varieties, the phi value is 0. This means that the lexical index between any given varieties lies between 1 and 0. In the Chinese dialects that Tang & van Heuven (2007) investigated, the phi value between any pair of varieties ranges from 0.698 to 0.079. These values were considered as the lexical distance among the selected Chinese varieties.

In other studies, lexical distance has been perceived in slightly different ways. For instance, in Van Bezooijen & Gooskens (2007), a study which was conducted on written Dutch, Frisian and Afrikaans, the lexical distance was considered as the percentage of non-cognate words in a sample of written text. The lexical distance between each pair of the three languages was determined based on the percentage of non-cognates, the percentage of words which are related via paradigm, and the lexical transparency. Words related via paradigm refers to words that may not necessarily genealogically related but just belong to the same paradigm. For instance, according to Van Bezooijen & Gooskens (2007), the Frisian *binne* (third person plural present tense of the verb 'to be') can be related to the Dutch *ben* (first person present tense of the verb to be) (see Van Bezooijen & Gooskens, 2007, p.255-257). Only a few studies are discussed here. There are other studies which employed lexicostatistics combined with functional and perceptual measures (e.g. Heeringa &

Nerbonne, 1999).

#### 4.2.2.5.3 Lexical Distances among Ethiosemitic Languages

Many studies have been conducted on Ethiosemitic languages using lexical comparisons (e.g. Bender, 1971; Bender & Cooper, 1976; Fleming, 1968 and Hudson, 2013). For example, Bender (1971) made lexical comparison among several Ethiopian languages. His study was a part of a nationwide project (from 1968 to 1969) which aimed at classifying the majority of (more than 80%) languages of Ethiopia into different proto-families. The study adopted 98-word lists from Swadesh (1950) basic vocabulary lists. The cognate identification was made based on the principle of ‘minimum one CVC’ correspondence. According to this parameter, a pair of basic vocabulary needs to share at least one vowel and two consonants to be considered as cognates. It seems that this method has its own shortcoming since, for example, Amharic words [t’əllələ] ‘purified’ and [k’əllələ] ‘became light’ have shared CVC correspondents, but they are not cognates; there is no historical relationship between the two words. In other words, Bender’s approach does not exclude false friends. Moreover, Bender (1971) did not explicitly mention how borrowed words were treated in the study. The basic cognate frequencies were computed using a computer and visual inspection. The computed frequency of cognates among thirteen Ethiosemitic languages treated in Bender (1971) is presented in Table 4.5. For the sake of specificity, only the results of the <sup>9</sup> Ethiosemitic languages are presented here (see Bender, 1971 for discussion on non-Semitic languages).

---

<sup>9</sup>AM = Amharic, AR = Argoba, HR = Harari, ZA = Zay, WE = Wolane, IN = Inor, CH = Chaha, MS = Mesqan, KS = Kistane, GA = Gyeto, GE = Ge’ez, TN = Tigrigna, TE = Tigre

Table 4.5: Percentage of cognates shared among 13 ES languages, Bender (1971)

	TE	TN	GE	GA	KS	MQ	CH	IN	WE	ZA	HR	AR
AM	52	56	62	65	66	56	55	50	60	59	61	79
AR	45	53	56	57	64	57	55	53	57	56	56	
HR	47	48	53	52	61	52	54	54	64	70		
ZA	49	47	54	53	61	55	58	55	79			
WE	52	51	61	59	63	59	62	55				
IN	43	43	46	49	66	70	81					
CH	44	43	51	52	70	80						
MQ	43	45	52	54	69							
KS	47	49	54	62								
GA	46	49	51									
GE	71	68										
TN	44											

Let us compare results in Table 4.5 with the classifications previously provided by historical linguists. For the sake of this illustration, the simplified version of the classification of Hetzron (1972) is presented in Figure 4.1. Table 4.5 shows that none of the East Gurage varieties (Harari, Zay and Wolane) is a dialect of another (based on 80% of shared cognate). The highest percentage of cognates is shared between Inor and Chaha (81%). Mesqan surprisingly shares 80% cognates with Chaha - the two language varieties are found in different groups in the classification of Hetzron (Figure 3.1). Eighty one percent (81%) cognate similarity between Chaha and Inor is also unexpected result since Inor is a Peripheral West Gurage language while Chaha is Central West Gurage language.

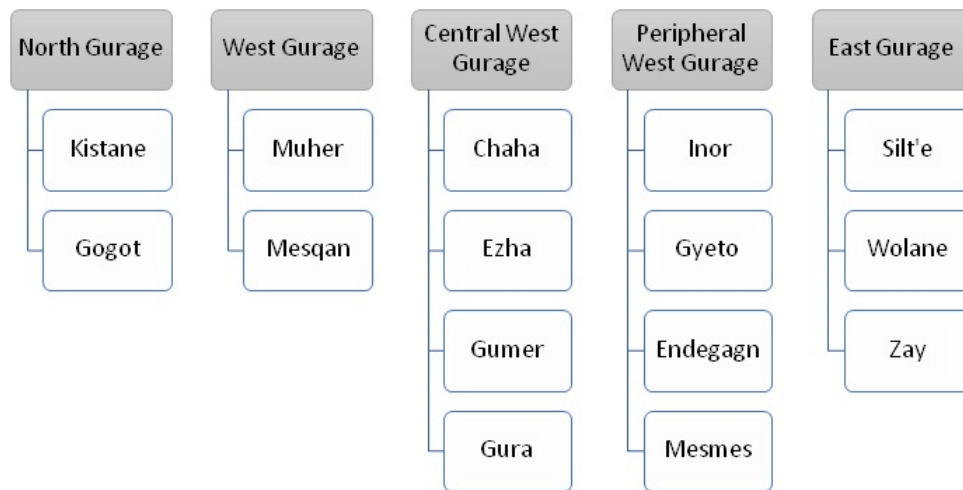


Figure 4.1: Summary of classification of South Ethiosemitic Languages

Like Bender (1971), Hudson (2013) compared fourteen (14) Ethiosemitic languages using 250 word list. The list of words were taken from Bender (1971), from lists of basic vocabulary (Swadesh, 1955) and from the Etymological Dictionary of Gurage, Volume II by Leslau (1979). The cognates were identified based on the form similarity among the words, combined with personal intuitive judgment. Similar to Bender (1971), borrowed words did not receive attentions in the vocabulary selections. Hudson (2013) argues that borrowed words are difficult to recognize especially when they are ‘nativized’ to the phonology and the meaning of the borrowing languages. Hence, he was convinced that the search for ‘the true cognate’ should be a serious issue only in a conservative study of Glottochronology when the aim of the study is speculating the date of the historical split between closely related languages (Hudson, 2013, p.63-65). Table 4.6 presents the frequency of the cognates in the basic vocabulary of the fourteen languages. Hudson (2013) reported just the counted data. They are converted to percentage<sup>10</sup> so that to compare them with results obtained from Bender (1971).

<sup>10</sup>In some languages, the counted percentage of cognate is more than 250 (104% and 103%) since Hudson considered synonyms as cognates.

Table 4.6: Percentage of cognates shared among 14 ES languages, Hudson (2013)

	TE	TN	GE	GA	KS	MS	MU	CH	IN	SI	ZA	HR	AR
AM	64	84	79	67	75	58	59	54	52	54	59	52	91
AR	56	72	68	67	72	57	57	54	52	55	61	56	
HA	48	50	55	47	57	54	52	50	48	66	68		
ZA	49	52	53	51	68	61	60	55	53	77			
SI	44	49	51	46	65	72	66	64	62				
IN	43	48	52	52	70	86	95	103					
CH	44	50	53	52	72	54	101						
MU	46	53	56	56	82	96							
MS	45	52	55	56	81								
KS	51	61	62	66									
GA	46	58	58										
GE	94	98											
TN	91												

Table 4.6<sup>11</sup> shows that the three East Gurage varieties: Harari, Zay and Silt'e share the highest number of cognates among themselves as compared to the cognates they share with other varieties. This supports the grouping proposal presented in Figure 4.1. It is also important to notice that the percentage of cognates shared between Zay and Harari in Hudson (2013) and in Bender (1971) are not quite different, 68% and 70% respectively. Other results do not confirm the grouping proposal of Hetzron (1972) and others which is presented in Figure 4.1. For instance, Inor shares 95% cognates with Muher, but Inor is a Peripheral West Gurage language while Muher is a West Gurage language. In the same manner, Muher shares 81% cognate with Kistane which is the North Gurage language. Likewise, the position of Mesqan seems controversial, it shares 81% cognate with Kistane. It is very important to notice that the percentages of cognates shared between Inor and Muher, Muher

<sup>11</sup>AM = Amharic, AR = Argoba, HA = Harari, ZA = Zay, SI = Silt'e, IN = Inor, CH, Chaha, MU = Muher, MS = Mesqan, KS = Kistane, GA = Gafat, GE = Ge'ez, TN = Tigrigna, TE = Tigre. In some pair of languages, the percentage exceeds 100% since Hudson considered synonyms as cognate words.

and Kistane and Mesqan and Kistane are larger than the percentages of cognates shared among the members of East Gurage varieties (Harari, Zay and Silt'e).

Furthermore, the high degree of similarity between Zay and Wolane (79%) and between Zay and Silt'e (77%) in Hudson (2013) show that Zay is very close to the two east Gurage languages - Wolane and Silt'e. The comparison between the results reported by Bender (1971) and Hudson (2013) also shows a great deal of disparity among the distances. For instance, Bender (1971) reported 80% lexical similarity between Chaha and Mesqan. However, Hudson (2013) reported just 54% lexical similarity between the two language varieties. Another significant difference is between Mesqan and Inor. While Bender (1971) reported 70% lexical similarity between the two varieties, Hudson (2013) reported more than 80% lexical similarity between the two. The same is true for the lexical similarity between Mesqan and Kistane. Bender (1971) reported 69% lexical similarity while Hudson (2013) reported 81% lexical similarity. This difference may be attributed to the sample size since Bender (1971) used 100 list of words while Hudson (2013) used 250 basic vocabulary lists. The approach of Hudson (2013) can also be questionable since he mixed cognates with synonym words.

Bender (1971) and Hudson (2013) are not the only studies that employed lexicostatistics to classify Ethiosemitic languages. Cohen (1961) also employed the same method to determine the distance among seven Ethiosemitic languages. He employed 116 basic vocabularies from Swadesh (1955) word list with some modification on the basic vocabulary lists. Regardless of a few number of languages included, the results supports the grouping of the North Ethiosemitic languages: Tigre, Tigrigna, and Ge'ez. However, a considerable variance in scores of the North against South Ethiosemitic languages was found, for example Tigre-Gafat 48% vs. Tigre-Harari 60% lexical similarity. The detail of the methods employed in this study is not presented here since only a few South Ethiosemitic languages were examined in the study. Comparison of frequency of basic cognates was also employed by Leyew & Siebert (2001) to determine the distance between Amharic and Argoba of Shewa-Robit and Aliyu Amba. They counted the number of cognates in a list of 100 basic words of the two languages and reported that the two languages share



about 80% of the basic vocabulary.

Recently, Menuta (2015) used 255 list of words to determine what he called ‘inherent intelligibility’ among six Gurage varieties <sup>12</sup> (Kistane, Chaha, Inor, Mesqan, Muher and Wolane). Menuta (2015) did not explain how the words were selected. He classified the words into three: completely similar words (if they have the same meaning and the same root consonant and vowel), partially similar words (have similar meaning and form, but differences in one or more of their sounds) and completely different words (have the same meaning but completely different form). Menuta (2015) does not talk about cognates. However, from his arguments, it seems that ‘completely similar’ and ‘partially similar’ words can be considered as cognates. Menuta (2015) presented each category separately. In Table 4.7, the mean of the completely similar words and those of partially similar words are added and their percentage (in terms of the total lexical item) are presented.

Table 4.7: Lexical similarity among six Gurage varieties, Menuta (2015)

	KS	CH	IN	MS	MU	WE
CH	62%					
IN	57%	85%				
MS	68%	87%	69%			
MU	70%	85%	75%	83%		
WE	47%	45%	42%	47%	47%	

According to Menuta (2015), Chaha and Inor, Chaha and Mesqan, and Chaha and Muher share more than 80% cognates. Mesqan and Muher are also very similar. The similarity between Chaha and Inor contradicts with the classifications previously presented by historical linguists since Chaha is a Central West Gurage language while Inor is a Peripheral West Gurage language. The same is true for the similarity between Chaha and Mesqan, and Chaha and Muher. Muher and Mesqan are West Gurage languages while Chaha is a Central West Gurage language. These again show a mismatch between the lexical distance and classifications by Hetzron (1972), Hetzron (1977) and by other historical linguists.

<sup>12</sup>KS = Kistane, CH = Chaha, IN = Inor, MS = Mesqan, MU = Muher, WE = Wolene

#### 4.2.2.6 Morphosyntactic Computational Approaches

The structural - computational approaches discussed in the preceding sections measure the distance among closely related languages based on either phonetic or lexical differences. The use of morphological and syntactic parameters has largely been only the focus of historical linguists who are rather interested in the classification of languages based on fortuitously selected specific features. The classifications of Ethiosemitic languages discussed in Chapter three is an illustration of this kind of approach. This approach has some obvious limitations. First, the parameters (the morphosyntactic features) are usually cherry-picked based on individual interest. Second, there is no clear threshold about the number of required shared innovations for varieties that are considered similar. The approach focuses on a few deliberately selected shared features while potentially there could exist several features that are skipped either intentionally or due to lack of adequate data.

There are a few recent studies that dealt with these limitations by applying computational approaches to morphosyntactic features. Spruit (2005) is one of them. The study was conducted based on reviews of syntactic variables in seven syntactic subdomains (complementizers, subject pronouns, expletives, subject doubling, subject cliticisation, reflexive and reciprocal pronouns, and fronting). A variable was defined as a linguistic unit in which two languages can vary. The syntactic distance was computed based on the frequency of syntactic variables in each of the seven syntactic subdomains. The dialect differences are aggregated to obtain a total syntactic distance. Then, the resulting map of the Dutch dialects with respect to syntactic variation was compared with the Daan and Block Dutch perceptual map. In later extended work, Spruit (2006) compared the geographical distribution of the syntactic variables in Dutch dialects with the map of Dutch dialects based on the measure of pronunciation differences (Heeringa, 2004). The geographical distances were also correlated with the syntactic distances using regression analysis.

Sampson (2000) and Nerbonne & Wiersma (2006) are another important contributions. Sampson (2000) determined the syntactic distance based on path differences in syntactic trees; the approach is called *leaf-ancestor paths*. Leaf-ancestor

path represents the syntactic structure by aggregating nodes starting from each leaf and proceeding up to the root. The leaves are parts of speeches while the ancestor is the highest mother node in the hierarchy of tree. Nerbonne & Wiersma (2006) used frequencies of trigrams of POS categories to detect syntactic differences between two varieties of Finish-Australian English; one spoken by the first generation and other by the second generation. The technique compares frequencies of trigrams of POS categories as indicators of syntactic distance between the varieties and then examines the potential effects of language contact. The frequency vectors were compared and analyzed by using permutation test. In more related work, Sanders (2007) combined part-of-speech trigram (see Nerbonne & Wiersma, 2006) and leaf-ancestors path (see Sampson, 2000) to determine dialectal variation in eleven British regions. Using International Corpus of English, the study illustrated that dialectal variation in selected British regions is detectable by using both leaf-ancestor paths and trigrams.

Heeringa et al. (2018) measured the syntactic distance among Scandinavian languages using a slightly different approach, *movement* and *indle*. Movement refers to average number of words that have been moved in a sentence of one language as compared to another language whereas indle is the average number of words being inserted or deleted in sentences of one language compared to the corresponding sentences in another language. They compared the results of these measures to the part of speech trigram measure (see Nerbonne & Wiersma, 2006) and reported a high correlation between the results of movement and that of the trigram. Nonetheless, the study reported a low correlation between the results of movement and that of indel, indicating that the two measures represent different kinds of linguistic variation. Gooskens & Heuven (2018) also examined the correlation between intelligibility among several European languages and syntactic (also phonetic, lexical, orthographic, morphological) distance among the languages. The study employed trigram frequencies (see Nerbonne & Wiersma, 2006) to measure the syntactic distance, and reported that among the linguistic variables, lexical distance is the best predictor of mutual intelligibility.

Heeringa et al. (2014) also investigated the impact of orthographic stems and

inflectional affixes on intelligibility among selected Germanic languages. The inflectional affixes were extracted from the stems and their differences were computed using Levenshtein algorithm. The same Levenshtein algorithm was used to compute the distance among the stems. The study reported very low correlation between orthographic stem variation among languages and orthographic variation in inflectional affixes. The study further reported that orthographic stem variation among languages is larger than orthographic variation in the inflectional affixes. It also indicated that orthographic lexical variation is strongly correlates with mutual intelligibility. Moreover, Hilton et al. (2013) examined whether Danes' comprehension of closely related language, Norwegian, is impeded by certain Norwegian grammatical constructions, using sentence comprehension. The study reported that word-order differences cause larger problems for listeners than morphological differences.

### **4.2.3 Functional Distance**

Functional distance refers to the difference between related languages, measured from the perspectives of the speakers' understanding or level of communication (see §1.2.5.2). The term 'functional distance' has been derived from 'functional intelligibility' which was used in Tang & van Heuven (2009). As indicated in 1.2.5.2, in the present study, functional distance means the distance among related languages which is measured based on degree of mutual intelligibility.

#### **4.2.3.1 Intelligibility and Interlingual Comprehension**

Mutual intelligibility is sometimes called 'functional measure' or 'test the informant' method (Tang & van Heuven, 2009, p.711). Mutual Intelligibility denotes the degree to which a language is understood by the speaker of another language because of the historical relationship among the languages (Gooskens et al., 2010, p.1-2; Gooskens, 2013. p.195; Gooskens, 2018, p.205). If the speakers of a language 'A' freely communicate with the speakers of language 'B' without having a direct exposure to language 'B', the two languages are said to be mutually intelligible (Gutt, 1980, p.57-58). Tang & van Heuven (2009, p.710) provided a quite similar definition. If two languages are mutually intelligible, they are also considered similar. The-

oretically, mutual intelligibility is associated with genealogical relationship among languages excluding contact-related factors (Gutt, 1980, p.57-58). Accordingly, mutual intelligibility (inherent intelligibility) and interlingual comprehension (acquired intelligibility) are different concepts. Interlingual comprehension refers to how much speakers of language ‘A’ understand speakers of language ‘B’ because of the exposure to both languages or the contact between the two languages (Gutt, 1980, p.58). Nevertheless, not all scholars agree on this division. For instance, Gooskens & Heuven (2018) treats both inherent and acquired intelligibility indiscriminately.

The term ‘mutual’ in mutual intelligibility implies that the direction of understanding between the speakers of two languages is two-way or bidirectional. However, in many studies (Ahland, 2003; Gooskens et al., 2010; Gutt, 1980; Golubović & Sokolić, 2013), it has been repeatedly attested that intelligibility is sometimes unidirectional. In other words, intelligibility is not always symmetrical. The speakers of language ‘A’ may understand well the speaker of language ‘B’, but not necessarily the vice versa, and this difference can be due to the characteristics of the languages (see Gooskens et al., 2010) and the influence of non-linguistic factors (Gooskens, 2007). Some languages can be relatively incomprehensible because of their complex linguistic features. For example, it has been reported that Danish speakers understand Swedish speakers much better than the Swedish speakers understand Danish speakers. The absence of symmetry in the intelligibility is often also linked to the absence of phonetic and prosodic features such as reduction phenomena including schwa assimilation and consonant reduction (Bleses et al., 2008, p.640). The degree of intelligibility is, therefore, sometimes defined as an average of the intelligibility of language ‘A’ for speakers ‘B’ and vice versa (Tang & van Heuven, 2009, p.722). However, this conceptualization poses epistemic challenges since ‘mutual intelligibility’ in principle is a symmetric concept.

Another important issue is the difference between intelligibility and language similarity. Some previous studies (e.g., Menuta, 2015; Tang & van Heuven, 2009) used lexical similarity to predict the degree of intelligibility among related languages. Though mutually intelligible languages are usually considered ‘similar’, this is not always the case. As indicated in 1.1, language similarity is often uni-dimensional;

two languages can be similar just on one dimension, for example, due to lexical similarity. Nonetheless, if two languages are only lexically similar, they are similar just from the lexical point of view. This does not mean they are morphologically or syntactically similar, too. Anyway, in principle, intelligibility is comprehensive, and contains all aspects of language features and even non-linguistic variables (Gooskens et al., 2010, p.10). Therefore, it appears to me that while mutually intelligible languages are probably always similar, the reverse claim is not always true.

When two languages are mutually intelligible, they are usually considered dialects of the same language (Gooskens, 2013, p.195; Gooskens, 2018, p.205; Menuta, 2015, p.86; Tang & van Heuven, 2009, p.207). If the intelligibility score drops below certain threshold (usually 70%- 80% score on intelligibility tests), the varieties are considered as different languages. In other words, intelligibility measures are employed as an alternative means of determining the distance among related languages. The general convention is that the more the languages are mutually intelligible, the more it is likely that they are similar to each other. Nonetheless, as indicated at the beginning of this section, the relationship between language similarity and mutual intelligibility is not always straightforward. The main challenge in this regard is that the degree of intelligibility among related languages is not always symmetrical. In other words, it may not precisely divide varieties into languages and dialects. Often studies use the mean of the two dimensions of intelligibility measures to solve the problem asymmetry. However, the degree of precision of using such a mean is an issue that requires further investigations.

The mismatch between language similarity and degree of intelligibility can also be due to the interference of extra-linguistic factors. There are cases when unintelligible varieties are considered dialect, e.g., Chinese dialects, and contrary to this, mutually intelligible varieties can also be considered as separate languages, e.g., Scandinavian languages. Such a mismatch is usually associated to political, social and historical situations. The lack of a precise method of measuring intelligibility can be another cause of mismatch. According to Tang & van Heuven (2009, p.712), intelligibility tests can be effectively implemented when they are applied to a small number of languages, but as the number of languages increases, it fails to become an effective

tool. This is largely because of the difficulty of controlling the priming effects that result from the repetition of the same materials across the speakers of several closely related languages.

Another challenge associated with using intelligibility as a measure of linguistic distance is the difficulty of determining the intelligibility threshold. There have been discrepancy across studies in the level of threshold they employ to determine whether a given languages are intelligible or not (see Casad, 1987). According to Grimes (1995, p.22), an intelligibility level of 85% constitutes dialect clusters of the same language. People with this level of understanding can communicate with each other without much difficulty. An intelligibility level between 70-85% shows that the linguistic groups can communicate, with certain degree of misunderstanding. Language varieties with intelligibility level below 70% are distinct languages. However, these suggestions emanate from intuitive judgments and different studies have been using different degree of threshold to determine the degree of intelligibility.

Different methods have been developed in the past decades to measure degree of intelligibility and interlingual comprehension among related languages (see Casad, 1987; Gooskens, 2013; Gooskens, 2018). Intelligibility tests can be broadly classified into two: opinion and functional (see Wang, 2007, p.25; Gooskens, 2013, p.197-208). While the functional test measures the level of understanding among the speakers of related languages, an opinion test measures the distance among related languages based on the subjective judgment of the speakers. The opinion tests will be discussed in 4.2.4. Intelligibility tests can also be classified into two based on the time of response provision: (1) online test (the test-takers respond while they are processing the stimuli); (2) offline test (allow subjects time to reflect before issuing the response) (Gooskens, 2013, p.201-205; Menuta, 2015, p.39; Wang, 2007, p.25). Intelligibility tests can also be categorized based on the structural complexity of the test items: word level, sentence level or text level which may include word recognition, sentence repetition and comprehension test (Gooskens, 2013, p.201-208; Menuta, 2015, p.163-167). Word recognition - involves recognizing a particular word at the end of a sentence. There are also different types of comprehension tests (Wang, 2007, p.30) including comprehension question (the test-takers answer questions after

reading/listening to a given input), sentence verification (listeners answer whether a statement is true or false), descriptive language (online comprehension question in which the participants are provided with visual presentations such as pictures or scenes from a movie or spoken description of pictures). Then, the participants determine if the description fits to the picture or scene by responding ‘correct’ or ‘wrong’. There are many other methods; some of them will be discussed in the next section.

#### **4.2.3.2 Intelligibility in Previous Studies**

According to Tang & van Heuven (2009, p.711), measuring mutual intelligibility was initially introduced by American structuralists Voegelin & Harris (1951). Since then, as noted above, various methods of measuring intelligibility have been developed (see Gooskens, 2013; Gooskens, 2018; Grimes, 1995; Tang & van Heuven, 2009; Wang, 2007). In this section, examples of different methods will be introduced. Tang & van Heuven (2009) employed two experimental methods to determine the intelligibility among Chinese dialects: a word Categorization test and word recognition in a sentence. In the Word Categorization test, listeners were asked to indicate to which of 10 given semantic categories a spoken word for example ‘apple’ belongs. The test-takers were expected to categorize ‘apple’ under ‘fruit’. The assumption here was that the categorization can only be achieved if the listeners correctly understand the target words. In word recognition in a sentence, the listeners were asked to write down words which could complete orally presented sentences. Each sentence was constructed in such a way that the final word could be predicted from the preceding part of the sentence.

Some other studies used translation as a tool to determine the intelligibility among related languages. For example, Beijering et al. (2008) employed the translation of the fable ‘The North Wind and the Sun’ to determine the intelligibility among eighteen Scandinavian varieties. The same method was previously employed by Beijering & Heeringa (2007). The sentences in the fable were translated from standard Danish to the 17 varieties. The participants then listened to the recorded sentences in each of the 17 varieties and re-wrote down in standard Dutch. The



intelligibility score was the percentage of correctly translated words. Van Bezooijen & Gooskens (2007) employed comprehension question to determine the degree of intelligibility among three West Germanic languages (Dutch, Afrikaans and Frisian) and three North Germanic languages: Swedish, Danish and Norwegian. The aim of the study was to examine if the three West Germanic languages are as intelligible as the three North Germanic languages. To assess the intelligibility, selected test-takers listened to the news about 'A kangaroo running on the streets of Copenhagen'. The text was translated from original Norwegian to Danish and Swedish and from Dutch into Frisian and Afrikaans. Then, the listeners answered comprehension questions which were derived from the story.

Filling in the blanks (cloze test) is another method that has been employed to measure mutual intelligibility. Van Bezooijen & Gooskens (2007), for instance, examined the intelligibility of written Afrikaans and Frisian for the Dutch native speakers using cloze test. They selected two short news articles. Five nouns five verbs, five adjectives and five adverbs were taken out from each article and were written alphabetically above each article. The locations of the word classes in the articles were replaced by blank spaces. Then, the two texts were translated from Dutch to Frisian and Afrikaans. Then, the participants were asked to fill in the blank space using the alphabetically listed words. A somehow similar method was employed by Renée & Van den Berg (1999). They employed contextual guessing of meaning of words in selected fragments. The fragments were taken from the descriptions of pictures made by some selected participants. In each fragment, a noun is included and the test-takers were asked to guess the meaning of the noun in each fragment. The expressions in the fragments were written in Dutch except the nouns for which the meaning should be predicted based on the given context.

Gooskens et al. (2010) employed three kinds of intelligibility tests to determine the asymmetry of the intelligibility between Danish and Swedish dialects. 'Read sentence' is one of the methods. Selected the participants were provided with semantically unpredictable (semantically anomalous) sentences such as 'He drank the wall', and the reading speed of the participants were determined based on these sentences. The second method was 'spontaneous speech'. Two spontaneous speech

tests were used. The first one was a picture description task in which the participants were provided with pictures and requested to describe the picture in detail. The second one was a map talk. In this case, a pair of participants who knew each other sat in separate rooms. One of the subjects of a pair described directions on the map orally and the second subject followed the direction. The participants were native Danish speakers and native Swedish speakers who were university students in Copenhagen. The map talk was performed in both Swedish and Danish languages for both language groups.

#### 4.2.3.3 Intelligibility among Ethiosemitic Languages

The degree of intelligibility among the South Ethiosemitic languages has not been seriously investigated. As indicated in 1.1, many of the claims about the intelligibility of these languages is based on the intuitions of the fieldworkers. For instance, all East Gurage languages except Zay have been considered intelligible to each other (Hetzron, 1972, p.2). There has also been a perception that all Central West Gurage languages are mutually intelligible. Hetzron (1972) also suggested that speakers of the Peripheral West Gurage languages do not have problems of understanding each other. He indicated that communication problem may exist only between the speakers of Central West Gurage and Peripheral West Gurage. He also proposed that North Gurage languages are mutually intelligible. The claim of intelligibility between Wolane and Silt'e and between Kistane and Gogot was stated in Demeke (2001), Hetzron (1972), Hetzron (1977) and others. However, these studies did not directly measure the intelligibility among the varieties. There may be many other such kinds of claims which are not relevant here.

There are some studies that directly investigated the intelligibility of Gurage varieties. To my knowledge, Gutt (1980) is the first serious study on the intelligibility of Ethiosemitic languages. This study examined the degree of mutual intelligibility among three Gurage varieties: Silt'e (East Gurage), Kistane (North Gurage) and Chaha (Central West Gurage). Listening comprehension questions derived from 'a short simple story' was orally presented to selected participants. Ten oral comprehension questions were administered to 10 persons of each variety. The test-takers

were asked to listen to the audio comprehension questions and to respond orally to each of the questions that followed the story. In addition to the oral comprehension questions, interviews were administered to assess the background of the test-takers - exposure to other varieties and language background of the test-takers' families. Gutt (1980) also tested the inter-lingual comprehension of three varieties (Wolane, Inor and Mesqan) using questionnaires. The questionnaires focused on the degree of contact among the speakers of the three varieties and on how much the speakers think they understand each others' variety. The result of this study indicates that the three languages: Silt'e, Kistane and Chaha are not intelligible to each other.

Ahland (2003) examined the degree of intelligibility among eleven Gurage varieties. Similar to Gutt (1980), oral comprehension questions (based on purposely selected stories) were used. Ten questions were presented to selected test-takers after each story. The results of the study are presented in Table 4.8 . The table shows, based on 80% intelligibility threshold, Chaha is intelligible to the speakers of Ezha (88%), Muher (85%) and Gumer (89%). In the same manner, Ezha is intelligible to the speakers of Gumer (87%). Gyeto is intelligible to the speakers of Inor (83%) and Endegagn (94%). Inor is also intelligible to Endegagn speakers (89%). Gumer is intelligible to Ezha and to Endegagn speakers, 89% and 88% score respectively. In the same manner, Endegagn is intelligible to Inor (90%). Mesmes is intelligible only to the speakers of Muher (97%). Mesqan is intelligible to the speakers of Chaha (92%), Ezha (90%), Muher (97%) and Gyeto (86%). Dobbis is intelligible to the speakers of Muher (85%). Moreover, Kistane is intelligible only to the speakers of Dobbis (90%).

Table 4.8: Intelligibility among Gurage varieties, Ahland (2003)

Languages <sup>a</sup>	CH	EZ	MU	GY	IN	GU	ED	MS	DO	KS
Chaha	90	88	85	86	67	89	63	48	39	30
Ezha	-	97	-	-	87	-	-	-	-	-
Muher	74	73	93	-	-	-	-	71	73	53
Gyeto	-	-	-	98	83	-	94	67	-	-
Inor	78	-	-	75	89	67	89	-	-	-
Gumer	-	89	-	-	88	97	-	-	-	-
Endegagn	-	-	-	77	90	-	93	-	-	-
Mesmes	-	-	97	-	-	-	78	-	-	-
Mesqan	92	90	85	86	-	-	-	92	89	63
Dobbi	-	-	54	-	-	-	-	61	96	76
Kistane	-	-	-	-	-	-	-	67	90	98

<sup>a</sup>CH = Chaha, EZ = Ezha, MU = Muher, GY = Gyeto, IN = Inor, GU =Gumer, EN = Endegagn, MS = Mesqan, DO = Dobbi, KS = Kistane. The column (abbreviated languages represent the test languages; the rows stand for the languages of the test takers)

From Table 4.8, it is important to notice that Mesqan is intelligible to many of the varieties except to Kistane (63%). Based on this, Hudson (2013, p.12) suggested that ‘Mesqan’ may refer to Muslim communities that speak mixtures of Gurage varieties. He also referred to a statement of Hetzron (1977) that some speakers of Inneqore<sup>13</sup> consider themselves as Mesqan. Asymmetry of intelligibility can also be observed from the table. For instance, Chaha speaker understand 74% of Muher while Muher speakers understand 85% of Chaha. Furthermore, Kistane speakers understand 76% of Dobbi while Dobbi speakers understand 90% of Kistane. In similar manner, Mesqan speakers understand 48% of Chaha while Chaha speakers understand 92% of Mesqan. Likewise, Gyeto speakers understand 67% of Mesqan while Mesqan speakers understand 86% of Gyeto. There is a large gap particularly between varieties in the last pair. This difference, to a certain extent, may be accounted to the attitude of Mesqan speakers. Ahland (2003, p.51) indicated that

<sup>13</sup>Inneqore is one of the dialects of Silt’e

“When we tested the people in the Mesqan area, many of them protested that some of the words are Sebat-bet Gurage language”. It might also be the case that all test-takers were not the speakers of Mesqan as a mother tongue; maybe the researcher erroneously interviewed non-Mesqan speakers who live in the Mesqan area. It is also very important to note that Chaha speakers understand Mesqan (92%) better than they understand their own languages (90%). This discrepancy could be because of problems related to material selection, quality of recordings or the participants’ sampling.

The comparison of these results with the classification of the languages presented in Figure 4.1 above shows that the two findings are somehow comparable to each other. The intelligibility test results seem to confirm the four Gurage groups presented in Figure 4.1. The two North Gurage languages: Kistane and Gogot (Dobbi) are somehow similar, 76% score on the intelligibility test. However, they are not mutually intelligible (with 80% threshold) as predicted by Demeke (2001) and Hetzron (1972). The three Central West Gurage varieties (Chaha, Ezha and Gumer) are intelligible to one another. Gura, another Central West Gurage variety, was not included in the Ahland’s study. Two of the Peripheral West Gurage varieties: Gyeto and Endegagn are not intelligible (77% score on the intelligibility test). The striking result here is the intelligibility between Chaha and Muher (85%) and Gumer and Inor (88%). Chaha is a Central West Gurage language while Muher is West Gurage language (see Figure 4.1). However, the two varieties are mutually intelligible. In the same manner, Gumer is Central West Gurage while Inor is Peripheral West Gurage, but the two varieties are intelligible to each other. Unfortunately, none of the East Gurage languages was included in the study of Ahland (2003).

In a recent study Menuta (2015) has also investigated inter-group communication among six Gurage varieties (Kistane, Mesqan, Inor, Chaha, Muher and Wolane). The aim of the study was to address various linguistic and non-linguistic issues related to the Gurage varieties. Determining the degree of intelligibility and inter-group comprehension was one of the objectives. Different tests were used to measure the degree of intelligibility and inter-group comprehension: word recognition (words in different parts of sentences were recognized by the participants), sentence repeti-

tion (the informants listened to varieties of sentence and wrote down exactly what they have heard), sentence verification (contains sentences that are always true, and the participants responded by saying true or false), instruction (the participants perform certain action based on orally given instructions) and comprehension questions. The participants responded to 95<sup>14</sup> questions, all tests combined (see Menuta, 2015, p.67-69). Twelve (12) participants were purposely selected (native speakers and residents of the local area) from each language area. Menuta (2015) reported the results of each test separately. In Table 4.9, the participants' responses to the 95 items are summed together and converted into percentages for the sake of easy of illustration. Table 4.9 presents the mean of all the tests averaged per language.

Table 4.9: Participants' scores on the intelligibility tests by Menuta (2015)

	KS <sup>a</sup>	CH	IN	MS	MU	WE
KS	92 (97%)	36 (38%)	36 (38%)	75 (79%)	50 (53%)	55 (58%)
CH	65 (68%)	92 (97%)	82 (87%)	84 (88%)	71 (74%)	53 (56%)
IN	55(57%)	88 (93%)	92 (97%)	79(83%)	51 (54%)	44 (47%)
MS	83 (88%)	68 (72%)	56(59%)	90 (95%)	71 (74%)	71 (75%)
MU	77 (81%)	90 (95%)	65 (69%)	89 (94%)	93 (97%)	62 (65%)
WE	59 (63%)	39 (41%)	34 (36%)	63 (67%)	42 (42%)	86(91%)

<sup>a</sup>KS = Kistane, CH = Chaha, IN = Inor, MU = Muher, WE = Wolane. The numbers outside the bracket show the participants' scores (out of 95) while numbers within the bracket show the percentage of the scores.

According to Menuta (2015), Chaha is intelligible to Inor speakers (87%) and to Mesqan speakers (88%). Inor is intelligible to Chaha (93%) and to Mesqan (83%). Moreover, Mesqan is intelligible to Kistane (88%). Muher is intelligible to Chaha (95%) and to Mesqan (94%). These findings contradicts the classification of Hetzron and others presented in Figure 4.1. Menuta (2015) is the most detailed study as far as the degree of intelligibility of the Gurage language varieties is concerned. There are, however, some concerns about the methods of Menuta (2015). First, the

<sup>14</sup>45 word recognition, 10 sentence repetition, 20 comprehension questions, 10 sentence verification and 10 instruction questions

same test items were repeatedly used across the participants from different language sites. Hence, it is highly probable that the participants' responses were unnecessarily exaggerated because of the priming effect. Besides, in sentence repetition test, the participants were asked to write down what they heard. It is not clear which writing system the participants used to write their responses since all the languages included in the study are not written languages. Menuta (2015) did not provide detailed description on the procedures of test administration. There have been attempts to develop an orthography for the Gurage varieties. However, as far as I know, it has not been taught in schools.

## 4.2.4 Perceptual Distance

In 1.2.4.4, it was indicated that nonspecialists' knowledge about their variety and other surrounding language varieties can be used as a means of determining linguistic distance among closely related languages. In this section, this issue is discussed further.

### 4.2.4.1 The Notion of Perceptual Distance

It was indicated in 1.2.5.4 that perceptual distance is the degree of difference between two or more closely related languages which is determined based on the perception of the ordinary people. Across literature, different terms are used to express this notion: *perceptual similarity* (Bradlow et al., 2010, p.930-931), *perceived similarity* (Gooskens, 2018, p.205; Gooskens & Heuven, 2018, p.5 and Tang & van Heuven, 2009, p.710), and *non-specialists' consciousness* (Preston, 1999, p.115; Mase, 1999, p.71; Preston, 1999, p.xxvi; Sibata, 1999, p.101). Perceptual distance assumes that the language competency of ordinary people is reliable for determining the linguistic distance among related languages (Long, 1999, p.177). In previous studies, there is often lack of clarity on the difference between the speakers' attitude towards a language and the perceptual distance. I hold the view that perceptual distance is a measure of degree of similarity/difference among related languages, based on speakers' personal judgment about the degree of similarity between their native languages and other related languages. On the other hand, language attitude is the

speakers' opinion, ideas or prejudices with respect to another language.

Distinguishing one language variety from another based on the perception of an ordinary people is as old as the Old Testament. An often cited biblical example is the use of the word *shibboleth* in order to distinguish the Ephraimites invaders from the inhabitants of Gilead (Book of Judges, chapter 12). It is indicated in the scripture that the 'sh' in *shibboleth* was pronounced as [ʃ] by the inhabitants of Gilead, but as [s] by the Ephraimites. Gilead used this pronunciation difference to identify Ephraimites invaders from his own people. As a discipline, according to (Preston, 1999b, p. 177), Perceptual Dialectology or Folk Linguistics emerged in 19th<sup>C</sup>, but it was extensively developed in 20<sup>th</sup> century. Though Perceptual Dialectology is now a well-recognized discipline, there are yet several issues that remained unclear. One of the issues in this regard is the fact that Perceptual Dialectology, as a discipline, lacks a sound theoretical foundation. Usually, its methods are derived from other disciplines such as Geographical Anthropology, Social Anthropology and Social Psychology. Moreover, linguists sometimes question the validity of perception-based linguistic distance (see Preston, 1999; Preston, 2017). While some studies have shown a mapping between the perceptual distance and production-based linguistic distance (e.g., Mase, 1964 in Preston, 2017) others did not (e.g., Grootaers, 1959 and Sibata, 1959 in Preston, 2017). The consciousness and unconscious cognitive mechanism underpinning the perception of the speakers is another area that requires further investigation (see Preston, 2010).

#### 4.2.4.2 Measuring Perceptual Distance

Methods that have been used to determine the perceptual distance can be broadly classified into two: those that use language input (audio sample) and those that do not use language inputs (Gooskens, 2013, p.201-202). From the late 20th<sup>C</sup> studies, the Little Arrow methods used by Grootaers (1959), Sibata (1959) and Mase (1964) cited in Preston (2018) are among studies that did not use language inputs. The Little Arrow method is also called 'degree-of-difference' (Long, 1999, p.177). In this method, the participants were asked where people speak similarly or differently. Then the arrows were drawn from the participants' site to each surrounding sites



identified as ‘the same’. Areas that are not connected by the arrows are considered areas of different dialects. Some of the recent studies used map ordering (e.g., Preston, 1993). In the study, selected participants were asked to order the maps of different USA states as 1 (the same), 2 (a little different), 3 (different) and 4 (unintelligibly different), relative to the participants’ home dialect. This study reported a match between perceptual boundary and production-based linguistic boundary. Tamasi (2003) also employed card categorization to examine the distinctiveness of regional varieties in USA. Selected participants were provided with cards on which 50 USA states are named, and they were asked to sort them into piles of dialect similarity. Later, hierarchical clustering was performed to group the states. The result of this study did not show a strong match between the traditional USA dialect boundaries and the perceptual boundaries.

Gould & White (1974) employed drawing-a-map method. Selected participants were provided with regional maps and asked to draw the linguistic map. Bucholtz et al., 2007 also used the map drawing task combined with questionnaires about various social and cultural factors. The participants were asked to draw a boundary around each part of California where they believe people speak differently, and label the area. Maps were also used by Pearce (2009). However, in this case, rather than drawing maps, the participants were asked to rank the maps based on the linguistic similarity/different of the areas on the map as compared to their home language. The above methods are based on mental mapping of the participants, and these methods of measuring linguistic distance were driven from cultural geographers (Preston, 2010, p.277).

Input-based methods employ language inputs (recorded speech) up on which the participants’ judgments are based (Long, 1999, p.184). Preston (1996) is one of the examples. In this study, a recorded voice was played to the participants, and the participants were asked to associate the recording with the map areas. Montgomery (2007) used the so-called ‘starburst’ method. He asked the participants from various cities in the north of England to identify voice samples from different parts of the country by marking on the map where they thought the voice was from. In the same manner, Boughton (2006) examined the diversity and leveling in the northern

urban French pronunciation. In this study, participants from Pays de la Loire region in the north west of France were asked to identify the speakers' background and determine whether they are from urban or rural area. This study shows that the participants' responses were partly correct except some errors associated with social and psychological factors such as stereotyping.

Both input-based and opinion-based approaches presented above are general in the sense that they do not focus on a specific linguistic feature. There are studies which focused on feature-specific approach, based on the assumption that speakers are sensitive to specific linguistic features of certain varieties (e.g. Labov, 2001; Plichta & Preston, 2005 and Niedzielski, 1999). For instance, in Niedzielski (1999), selected participants (students of the Detroit area) took test in which they were asked to choose from a set of synchronized vowels the tokens that they felt matched best the vowels they heard in the speech of a fellow Detroiter. Some of the participants were told that the speaker is a Detroiter, and others were told that the speaker is a Canadian. According to this study, participants that were given the Canadian label chose raised-diphthong tokens as those present in the dialect of the speaker, whereas those that were given a Detroiter label did not. This result indicated how the participants' linguistic judgment is affected by stereotyping. Perception-based linguistic distance is usually classified into two: perceived similarity and perceived intelligibility.

#### 4.2.4.2.1 Perceived Similarity

Perceived similarity is non-linguists' judgment about the *similarity* between the speakers native language and other close varieties. Perceived similarity is also called *judged similarity* in Beijering et al. (2008, p.13), and *perceived distance* in Tang & van Heuven (2009, p.710) and Gooskens (1997, p.5). Many studies previously employed perceived similarity to determine the distance among closely related languages. For instance, Gooskens (1997) used judged similarity to determine the distance between the standard Dutch and other Dutch dialects. Some selected speakers of standard Dutch dialects listened to series of fragments and rated the similarity

against the standard Dutch on the scale that ranges between 1 and 10, where 1 is the variety in question and 10 is standard Dutch. The study reported that the perception of the speakers can be employed to determine the distance among closely related languages.

Gooskens (2005a) employed the judged similarity measure in order to determine to what extent the Norwegian speakers understand their own dialects. The audio inputs of the Norwegian dialects were provided in two forms: with and without intonation (manipulated speech signal). The two forms were included to examine the effect of suprasegmental features on the judgment of the listeners. The recorded readings (original and monotonized) were presented orally to the selected subjects who were the speakers of 15 Norwegian dialects. One of the findings of this study was that intonation is important in determining the distance among the Norwegian dialects. In the same manner, Beijering et al. (2008) measured perceived similarity to determine the distance among 18 Scandinavian varieties. Selected listeners (speakers of standard Danish) were classified into three groups, each group responded to the recordings of six varieties; the 18 varieties are distributed among the three groups to minimize the priming effect. The listeners responded on the scale that ranges from 0 (similar to the standard Danish) and 10 (not similar to standard Danish). Gooskens & Heeringa (2004) used a similar method to determine the perceptual distance among Norwegian dialects. The listeners were asked to judge the recordings on a scale from 1 (similar to one's own dialect) to 10 (different from one's own dialect).

#### 4.2.4.2.2 Perceived Intelligibility

Perceived intelligibility is a measure of the degree of *intelligibility* between the speakers' native language and other local varieties, based on the perception of non-linguists. In some studies it is called *judged intelligibility* (e.g. Tang & van Heuven, 2009, p.709). The methods of measuring perceived intelligibility is usually very similar to the methods of measuring perceived similarity except that in perceived intelligibility tests participants are asked about the degree of intelligibility between their native language and other related languages. For instance, in Tang & van

Heuven (2009), the participants were asked to rate how much they understand a given Chinese dialect on the Likert scale that ranges from 0 to 10.

In the same manner, Tang & van Heuven (2007) employed a recording of ‘The North wind and the Sun’ to determine the perceived intelligibility among 15 Chinese dialects. Two versions of the fable were produced; one with original pitch interval kept intact, and the other one with all the pitch movements replaced by constant pitch (monotone). The monotonized version was administered to examine the effect of tone on the intelligibility judgment. The test was administered to 360 participants, 24 participants from 15 dialect areas. The participants rated the recordings of the fable on a scale from 0 (do not understand a single word of the speaker) to 10 (understand the speaker perfectly). There are many other recent studies that examined the perceived intelligibility among related languages (e.g., Fossen, 2018; Gooskens & Heuven, 2018; Kürschner et al., 2008). In general terms, these studies seem to suggest that the linguistic knowledge of ordinary people can be used to determine the degree of intelligibility among closely related languages. Gooskens (2013) seem to suggest that production based perceptual distance is more reliable as an indicator of linguistic distance than perceptual distance which is measured without language input.

### **4.3 Relationship among the Three Dimensions**

In section 1.1.1 it was indicated that examining the degree of relationship among the three dimensions of distance is one of the main objectives of the present study. Furthermore, in 4.2.4 it turned out that, in previous studies, there have been attempts to determine to what extent the perceptual dialect area corresponds to the production-based traditional dialect area. In the section that follows, studies which tried to address this issue are summarized.

#### **4.3.1 The Importance of Examining the Relationship**

It was indicated in 1.1 that precisely quantifying the linguistic distance is not easy. The main reason is that each method of measuring linguistic distance has its own

drawbacks. Structural distance is usually unidimensional ,i.e., it measures linguistic distance either from phonetic or lexical point of view while the similarity/differences among languages is usually multidimensional. Functional distance is usually influenced by several non-linguistic factors (see Gooskens, 2013; Gooskens et al., 2010; Gooskens, 2018). Hence, there is no a perfect way of determining linguistic distance. A possible way of facing this challenge is comparing results across different distance measures and checking the correlations among the results. Strong correlation among all the measures can be taken as confirmation of the similarity or difference among the languages. As also reported in the previous studies (e.g., Gooskens et al., 2010; Gooskens & Heuven, 2018; Tang & van Heuven, 2007; Tang & van Heuven, 2009), the correlations results can be used to determine, for instance, whether the structure-based language similarity is a sufficient requirement for the the languages to be mutually intelligible, or whether mutually intelligible languages are always similar.

The correlation results can also be used to examine the degree of similarity between the structure-based linguistic distance and the distance based on the perception of non-linguists. This in turn is very important to address issues related to the validity of the perceptual distance which was discussed in 4.2.4. Furthermore, determining to what extent the three dimensions of distance are related is essential for the decision someone has to make in order to choose the best approach while measuring the distance among related languages. Moreover, strong correlations among the three dimensions may give a good foundation to combine different methods and examine linguistic distance from different perspectives. Related to this, Misao (1953, p.11) cited in Preston (2018, p.181) suggests that “work on regional speech should go forward only after folk ideas about languages [were] determined”. It was also indicated in 4.2.3.1 that designing intelligibility tests for many closely related language is difficult. Strong correlation among the three dimensions of distance means one of the dimensions of distance can be sufficient to make a valid claim about the distance among related languages as far as the appropriate distance is carefully chosen.

### 4.3.2 Approaches towards Examining the Relationship

Several studies have previously examined the relationship among the three dimensions of distance (e.g., Beijering et al., 2008; Gooskens & Heeringa, 2004; Gooskens & Van Bezooijen, 2006; Tang & van Heuven, 2007; Tang & van Heuven, 2009; Tang & Van Heuven, 2015; Tang & van Heuven, 2015). Among these studies, many of them reported a positive correlation. However, the results are not consistent across all the studies. For instance, the correlations between the structural distance and the functional distance reported in Tang & van Heuven (2009) and in Van Bezooijen & Gooskens (2007) contradict each other. In Tang & van Heuven (2009), no significant correlation was found between the phonetic distance and functional distance. However, the result reported in Van Bezooijen & Gooskens (2007) is the perfect opposite, i.e., the study reported a strong correlation between phonetic distance and functional distance. Likewise, a strong correlation between lexical distance and mutual intelligibility was reported in Tang & van Heuven (2015) while no strong correlation between phonetic distance and degree of intelligibility was found.

Gooskens (2018, p.209) argues that lexical distance predicts the degree of intelligibility among related languages only to a certain extent (based on studies conducted on the Scandinavian languages). She also provided several explanations for this. One of her arguments is that a single lexical item may have unpredictable impact on the degree of intelligibility since a non-cognate lexical item can obscure the understanding of the whole sentence or speech, especially if the non-cognate word carries the central meaning of the whole message. Moreover, it can also be the case that learners may easily understand non-cognate words if the words were previously borrowed into the native languages of the learners. False friends may also cause a huge problem since they mislead the learners. According to Gooskens (2018, p.210) phonetic distance is a good predictor of mutual intelligibility, at least in the context of Scandinavian languages. She indicated that some sounds are more important than others in terms of determining the degree of intelligibility - for example, consonants are more important than vowels since they are 'the framework' of the words. She also further indicated that word recognition is the key to speech

understanding; if a speaker precisely recognizes minimal proportion of words, he/she can easily decode the messages in a speech. Tang & van Heuven (2007) seems to substantiate this argument. In this study, a strong correlation was reported between phonetic distance and mutual intelligibility whereas there was no strong correlation between the lexical distance and mutual intelligibility.

## 4.4 The Determinants of Linguistic Distance

In this section, some of the factors that determine the linguistic distance among related languages are discussed. For the sake of convenience, the determinants of the linguistic distance are classified into two: linguistic and non-linguistic determinants.

### 4.4.1 Linguistic Determinants

There may be several linguistic factors that determine the distance among related languages. The present study focuses on three of them; characteristic features, shared features due to contact and the influence of local dominant languages.

#### 4.4.1.1 Characteristic Features

Related languages always have some shared common features, whether these features are phonetic, lexical or grammatical. In the traditional structural approaches, for example in comparative linguistics and dialectology, there are always specific linguistic features that are accountable for the differences and similarities among related languages. This section aims to provide the summary of techniques of detecting characteristic features and extracting them from the aggregate data. Scholars have recently invented computational means of determining phonetic features that characterizes closely related languages (Nerbonne et al., 2011, p.83-85; Prokić et al., 2012 and Wieling & Nerbonne, 2011).

For instance, Wieling & Nerbonne (2011) presents two quantitative means of measuring the degree to which a feature identifies a dialect area: *representativeness* and *distinctiveness*. A feature in a dialect area is distinctive with respect to the larger language area to the degree that it occurs exclusively in that area. Representitiveness

refers to the degree at which a feature is found throughout the sites in a group. In other words, a feature increases in representativeness to the degree that it is found at each site in the group. Since the mathematical conception behind distinctiveness and representativeness is not the focus of the present study (for detail see Prokić et al., 2012, p.72-74 and Wieling & Nerbonne, 2011), only the simplified version of representativeness is presented below. Given  $f$  the feature in question,  $g$  the set of sites in the given cluster and  $g^f$  the set of sites where feature  $f$  is observed, the representativeness of the features can be denoted as follow.

$$Representativeness(f, g) = \frac{|g^f|}{|g|}$$

According to Prokić et al. (2012, p.73), the concept of distinctiveness and representativeness has two limitations; first, the formulation applies only to individual value of categorical features, not to the features themselves. Second, on the contrary, many dialectological analyses are based on numerical measures of feature differences, for instance the edit distance between two pronunciations. As a result, Prokić et al. (2012) presented an alternative approach that can be applied to the feature difference, not just to their categorical values. The alternative approach involves finding features that are within the cluster of interest while different from the rest of the sites. The two components, i.e., *within difference* and *between difference* characterize the average difference within the cluster (within difference) and the average difference between the sites in the cluster and the rest of the sites in the data set with respect to given item (between difference). Features whose within difference is low and between difference is high are regarded as characteristic features or *shibboleths*.

Figure 4.2 shows one of such characteristic features of the Central West Gurage languages (Chaha, Ezha, Gura and Gumer) based on the pronunciation of the word 'traveler'. The word is pronounced as either [əməṇə] or [əmənnə] in the Central West Gurage area (illustrated by dark blue color) while it is pronounced as either [iməñña] or [məññə] in the remaining parts of the Gurage area. In this particular example, [n] is the phonetic feature within cluster of interest (Central West Gurage) that makes the Central West Gurage dialect area distinct from other Gurage areas. [n] is pronounced as [ñ] outside the Central West Gurage area.



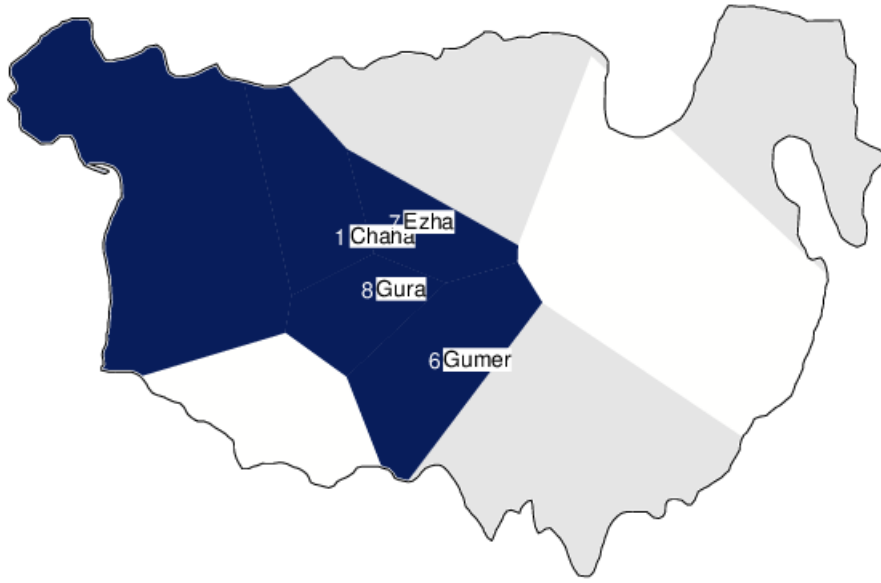


Figure 4.2: Distribution map: 'traveler' selected as one of the characteristic feature of the Central West Gurage

The basic notion of within difference and between differences can be summarized as follow. Given  $g$  = a group of languages to be examined,  $|g|$  = the number of sites of the group of languages to be examined,  $G$  = a large area of interest,  $|G|$  the number of sites including the sites  $s$  both within and outside  $g$ ,  $d$  = measure of differences between sites with respect to the given feature  $f$ , and  $d^f$  = the set of sites where feature  $f$  is observed, the mean difference with respect to  $f$  within the group is calculated using the following formula.

$$d_f^{\bar{g}} = \frac{2}{|g|^2 - |g|} \sum_{s, s' \in g} d^f(s, s')$$

The mean difference with respect to  $f$  involving elements from outside the group is computed using the following formula.

$$d_f^{\bar{g}'} = \frac{1}{|g|(|G| - |g|)} \sum_{s \in g, s' \notin g} d_f(s, s')$$

As indicated above, the characteristic features were considered as relatively large difference between

$$d_f^{\bar{g}'} \quad \text{and} \quad d_f^{\bar{g}}$$

However, since the raw difference values differ based on the data sets, the above notations are usually normalized as follow so that the differences are comparable

across different data sets.

$$\frac{\bar{d}_f^{g'} - \bar{d}_f}{sd(d_f)} - \frac{\bar{d}_f^{\bar{g}} - \bar{d}_f}{sd(d_f)}$$

Gabmap uses these formulas to identify the characteristic features with relatively large differences between  $\bar{d}_f^{g'}$  and  $\bar{d}_f^{\bar{g}}$ . In Gabmap, linguistic features (e.g., pronunciation differences) that have large differences are ranked on the top of the input features and those that have low difference are ranked at the bottom of the list.

Some previous studies employed this method to identify specific features in groups of dialects. For instance, Prokić et al. (2012) employed this approach to determine the characteristic feature of six Dutch dialects: Frisian, Low Saxon, Franconian, Limburg, West Flanders, and Belg.Brabant. In the same manner, Prokić et al. (2013), employed the same approach to identify specific features of Bulgarian dialects. Wieling & Nerbonne (2011) used the method to determine the characteristic features of Dutch dialects. To add more, Bloem et al. (2016) used this approach to determine accent difference between English L1 and L2 speakers.

#### 4.4.1.2 Shared Features Due to Contact

In 4.2.2.5.1, it was indicated that language changes can be either because of horizontal diffusion (contact) or due to vertical relationship (genealogical relatedness). The phonetic and lexical distance presented 4.2.2.4 and 4.2.2.5 respectively do not explain whether the distance among the languages is due to contact or due to genealogical inheritance. This section presents two approaches that are usually used to determine whether the similarity or difference among related languages is because of contact or the result of genealogical relationship among the languages.

##### 4.4.1.2.1 Estimating Horizontal Diffusion

As noted above, approaches used for areal classifications, do not make a distinction between vertical and horizontal diffusion of features; hierarchical clustering makes no assumptions about the source of features, always grouping similar items together regardless of whether the source of their overlapping features was inheri-

tance from parent nodes or borrowing from sibling nodes [Nerbonne in discussion] Prokić et al. (2013, p. 151-153) further suggests that using hierarchical clustering such as the ones discussed in the previous sections can be sometimes problematic since usually we need to know whether the cause of the similarity between languages is diffusion of horizontal features or historically inherited features. Prokić et al. (2013, p.152) suggests using Neighbor-net algorithm to determine whether the classification of languages is really genealogical or areal. One of the properties of Neighbor-net algorithm is that if the input distances reflect a horizontal relationship, it gives a net-like structure. If the input distances are additive, it gives the corresponding binary tree. This properties are usually used to determine whether the groupings of the related languages are due to contact or genealogical relatedness (see Bower, 2012 for the basic conception underlying Neighbor-net algorithm). In the present study, Neighbor-net representations of the lexical distance was computed using Splits Tree (see Huson & Bryant, 2010).

#### 4.4.1.2.2 Estimating Borrowed Features

As indicated in 4.1, the Tree model is represents only in the vertical transmission of linguistic features. As a result, the historical linguists who are proponents of the Tree model have always been concerned about means of detecting and excluding features that are shared among languages because of contact. There are different methods of detecting borrowing; though their precision is often criticized (see Heine, 1974; Minett & Wang, 2003; Wang & Minett, 2005). One of the studies that is worth mentioning is Hinnebusch (1996). He proposed a lexicostatistical approach for identifying languages that come into contact by means of lexical skewing. By 'skewing' he means a measure of the asymmetry of phonetic or lexical distance values. The skewness value can be either positive or negative. In the words of Hinnebusch (1996, p.186) and Wang & Minett (2005, p.128), skewing is the difference that is observed between the similarity of one language with respect to two other languages. The basic assumption of lexicostatistical skewing is that languages which group together will tend to have a numerical symmetry with other related languages in the com-

parison set if indeed the grouped languages form genetic group (Hinnebusch, 1999, p.184; Wang & Minett, 2005, p.128). In other words, languages within a subfamily are likely to exhibit little skewing with respect to languages of another subfamily. If skewing is observed between sister languages, however, language contact is one possible cause.

According to Minett & Wang (2003, p.122-124) lexicostatistical skewing is the most effective way of detecting borrowing among genetically related languages. However, according to Wang & Minett (2005, p.273) it has also its own drawback since it does not consider retention of features because of, for example, the conservative behavior of a language. Additionally, it may not explain the contact between sister languages in a subgroup and other unrelated neighboring languages. Indeed, every computational method of detecting borrowing has its own limitation (see Cavalli-Sforza et al., 1994; Sankoff, 1972; Embleton, 1982). The present study employed lexical skewing to determine the percentage of borrowed phonetic and lexical features among the South Ethiosemitic languages since lexical skewing is particularly effective when it is applied to lexicostatistical data (Minett & Wang, 2003, p.123). Since this is the case in the present study, it was anticipated that the percentage of borrowed phonetic and lexical features can be predicted with high degree of accuracy.

Hinnebusch (1996) employed this basic idea of lexicostatistical skewing to identify lexical borrowing among several sub-families of Bantu languages. Accordingly, for example, Samburu and Nandi share 9.9% of the lexical items in 200 word basic vocabulary. However, Masai and Nandi share 15.7% of the basic vocabulary. It was assumed that the 5.8% difference was due to contact between Masai and Nandi. The notion of lexicostatistical skewing proposed by Hinnebusch (1996) is further illustrated below using Figure 4.2.

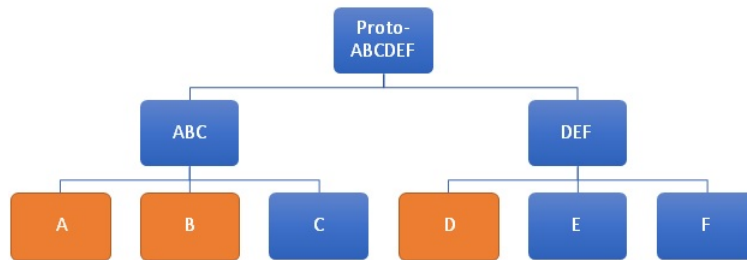


Figure 4.3: A diagram illustrating the nation of lexicostatistical skewing

In Figure 4.3, A, B, C are sister languages while D, E, F are also sister languages in the sister subfamily. The basic notion of lexicostatistical skewing is that any two sister languages, let say language 'A' and language 'B' must have symmetrical relationship with any of the language in other subfamily, let say language 'D'. Therefore, given two sister languages  $L_A$ ,  $L_B$  and a third language from another subfamily  $L_D$  and given the lexical similarity ( $S$ ) among each of these languages, the lexicostaistical skewing between  $L_A$  and  $L_B$  with respect to  $L_D$  can be denoted as follow.

$$S_{AD} - S_{BD}$$

If a contact has occurred between, let say, recipient language  $L_A$  and donor language  $L_D$ ,  $L_A$  will tend to be more similarity to  $L_D$  than  $L_B$ , resulting in a positive skewing between  $L_A$  and  $L_B$  with respect to  $L_D$ . This tendency of contact to have a positive skewing forms the base for skewing method (Wang & Minett, 2005, p.127).

Wang & Minett (2005, p.128) defined aggregate skewing of language  $L_A$  with respect to another language  $L_D$  as the average skewing between  $L_A$  and each of its siblings ( $L_B$  and  $L_C$ ) with respect to  $L_D$ . The direct interpretation of the positive aggregate skewing of  $L_A$  with respect to  $L_D$  is that  $L_A$  has more contact with  $L_D$  than

its siblings ( $L_B$  and  $L_C$ ) do. Negative skewing can have the opposite interpretation. It could mean that, compared to  $L_B$  and  $L_C$ ,  $L_A$  has less contact with  $L_D$ . Indeed many other interpretations are also plausible. For instance, it could be interpreted that (1)  $L_A$  is less resistant to change compared to  $L_B$  and  $L_C$ ; or (2)  $L_A$  may have contact with other languages outside its own subfamily. If there was no contact the skewing between language 'A' and language 'B' should have been zero. Method of computing lexicostatistical skewing is illustrated in Table 4.11 using sample lexical similarity matrix of Central West Gurage and West Gurage (Muher and Mesqan) language varieties presented in Table 4.10.

Table 4.10: Sample lexical similarity matrix

		CWG			WG		
Language <sup>a</sup>		CH	GM	GU	EZ	MS	MU
	Chaha	100					
CWG	Gumer	88	100				
	Gura	95	93	100			
	Ezha	92	92	90	100		
	Mesqan	89	87	88	88	100	
WG	Muher	86	85	85	87	88	100

<sup>a</sup>The test languages are abbreviated - EZ = Ezha, GM = Gurmer, GU = Gura, MS = Mesqan, MU = Muher, and CH = Chaha; the resulted are converted to percentage

Table 4.11 illustrates how aggregate skewing is computed. The aggregate skewing is computed for the Central West Gurage language *Chaha* with respect to the West Gurage varieties.

Table 4.11: Illustration of aggregate skewing of Chaha with respect to the West Gurage varieties

		Central Gurage	
%		Mesqan	Muher
	Gumer	88-87 = 1	86-85 = 1
	Gura	88-88 = 0	86-85 = 1
	Ezha	88-88 = 0	86-87 = -1
$\delta S_{CentralWestGurage}^{WestGurage}$	Chaha	= 1/3	= 1/3

Aggregate skewing of the other language varieties can be computed using the same procedure. In the table, a small positive aggregate skewing of Chaha (1/3%) with respect to Muher shows that Chaha has slightly more contact with Muher as compared to its sister Central West Gurage languages (Gura, Gumer and Ezha). Similar interpretation holds true for the skewing between Chaha and Mesqan.

One may raise questions regarding what amount of aggregate skewing is a significant indicator of contact. With regard to this, Wang & Minett (2005, p.131-135) examined the distribution of aggregate skewing in two conditions (*no contact* vs. *contact*) using Monte-Carlo simulation. For the 'no contact' condition, Pseudo-random character data are generated for ten languages that have splits into two sub-groups, each subgroup consisting of five languages. The distribution of aggregate skewing by Monte-Carlo simulation was estimated by observing the relative frequency of different values of aggregate skewing in multiple runs (more than 1000) of the algorithm. They found aggregate skewing that is very similar with Gaussian distribution, with almost zero mean, -.01. In the contact condition, borrowing of various degrees was injected between a single donor language and a single recipient language, one language in each subgroup. The result of this experiment indicated that the observed mean level of aggregate skewing of the recipient language with respect to the donor language is +3.3%. Based on these experiments, Wang & Minett (2005) came to the conclusion that, for 5% probability of false alarm (the significance), language contact can be inferred whenever the aggregate skewing exceeds the threshold value of 3.7%. Since this threshold may vary depending on the retention rates of languages of interest, in the present study, 4% and more aggregate skewing was assumed as a significant indicator of contact.

#### 4.4.1.3 The Influence of Dominant Local Languages

In section 2.2, it was indicated that Ethiopian linguistic area is often considered as a melting pot where languages from four different families (Semitic, Cushitic, Omotic and Nilo-Saharan) are intermingled through a long history of contact. It was further illustrated that several scholars, for example, Crass & Meyer (2009), Ferguson (1976), Tosco (2000) and Zaborski (1991) investigated the linguistic features of the

Ethiopian linguistic area. The implication of these studies is that it is not possible to discuss the distance among South Ethiosemitic languages without, at least, recognizing the influences of language contact. Hence, this section acknowledges some of the studies that directly dealt with the contact among Ethiosemitic languages and between Ethiosemitic and non-Semitic languages.

Although the contact situations in the area require further investigation, there are several contributions that are worth mentioning. For instance, Meyer (2005b) investigated language and cultural contact among three Gurage varieties: Muher, Zay and Wolane. The study examined cultural and linguistic convergence among the speakers of the three languages based on descriptions of selected riddles. Leslau (1956) also investigated Arabic loan words in the Gurage varieties. The study found out that these loan words are imported through merchants and the expansion of Islam. The greatest number of loan words were found in the Eastern Gurage languages and in Masqan-Gogot whose speakers are predominantly muslims. Compared to other Gurage varieties investigated by the study, the number of loaned words in Harari is very high. Furthermore, Leslau (1952) studied the influence of Sidamo on the Gurage varieties. The study reported several linguistic instances that do not exist in Semitic languages but exist in the Gurage varieties and Sidamo. In the same manner, Appleyard (2015) studied the contact between Semitic and Cushitic languages based on ancient contact features in Ge'ez and Amharic.

Crass & Meyer (2009) studied the morpho-syntactic and pragmatic functions of the quotative verbs in selected Ethiosemitic languages (Muher and Amharic), and compared them with quotative verbs in Low Land East Cushitic, Oromo. Moreover, Leslau (1945) studied the influence of Cushitic on Semitic languages. The study investigated features that exist in Ethiosemitic languages and in Cushitic languages but do not exist in the remaining Semitic languages, and listed several phonological, morphological and syntactic features that are borrowed from Cushitic to Ethiosemitic languages. It is also important to mention that Little (1974) discusses syntactic evidence of contact between Cushitic and Semitic.

Given these, the present study examines the influence of two dominant local languages, Amharic and Oromo, on the Gurage varieties based on phonetic and lexical



data. These two languages are chosen for a couple of reasons. First, the Gurage area is surrounded by Oromo speakers (see §2.1); hence, due to the geographical proximity, Oromo has a potential to influence the Gurage varieties. Second, Amharic has been the most influential language in the Gurage area since it has been the medium of school and administration for a century (see §2.2.2). Third, Amharic and Oromo have a large number of speakers, compared to all other languages spoken in Ethiopia. Hence, given the Gravity model (Trudgill, 1974) (also see §4.4.2.2), it is very likely that these two languages influence the Gurage varieties that have a small number of native speakers.

#### **4.4.2 Non-linguistic Determinants**

This section presents major non-linguistic variables that determine the linguistic distance. Though several non-linguistic variables may affect the linguistic distance, the focus here is only on the contact among the speakers, geographical distance and population size, and language attitude.

##### **4.4.2.1 Contact among the Speakers**

Previous studies show that contact among the speakers is the most influential factor that affect the distance among related languages. For instance, Gooskens & Heuven (2018) examined the influence of linguistic (phonology, orthographic, lexical, syntactic) and non-linguistic (exposure and attitude) on the intelligibility between 16 closely related European languages. The result of a step-wise regression analysis indicated that exposure is the main factors that determine the intelligibility among the languages. Similarly, in the study conducted on the similarity between Amharic and two Tigrigna varieties, Ethiopian and Eritrean Tigrigna, Feleke (2017) reported the contact among the speakers as one of the factors that contribute to the distance among the languages. In the same manner, Gooskens & Van Bezooijen (2006) reported the contribution of contact among the speakers in the study that examined the mutual intelligibility between written Frisian, Dutch and Afrikaans. In several other studies, for example in Schneider (2003), Trudgill (1983) and Trudgill (2004),

the contribution of language contact to languages change has been extensively discussed.

There are also studies that reported geographical issues such as river density and landscape roughness that are associated with contact among speakers affect the linguistic distance among related languages. For instance, Axelsen & Manrubia (2014) investigated the significance of 14 variables: landscape roughness, altitude, river density, distance to lakes, seasonal maximum, average and minimum temperature, precipitation and vegetation, and population density. This study indicated about 80% linguistic diversity in Africa is attributed to river density and landscape roughness. The contribution of these factors in other continents is relatively low; Australia and Pacific 69%, Europe 56%, and America 53%.

It is not always the case that the influence is one directional; there is also a tendency that linguistic distance affects the level of contact. According to a study conducted by Hutchinson (2005), countries that have a significant linguistic difference have less trade relationship as compared to countries that speak closely related languages. Since there are several studies (e.g. Gooskens, 2005; Gooskens & Van Bezooijen, 2006 Heeringa et al., 2011; Menuta, 2015 Nerbonne et al., 2005) that strongly confirm the influence of contact on linguistic distance, it will be considered as a potential factor that may also contribute to the distance among South Ethiosemitic languages.

#### **4.4.2.2 Geographical Distance and Population Size**

The Wave model introduced in 4.1 assumes that linguistic features spreads from a center like a concentric wave. The adjacent speech communities are always affected before the remote one (see Boberg, 2000; François, 2015; Heggarty, 2010). The features spread until they are blocked by some physical or social barrier or until their force is exhausted by geographical distance. The spread of innovations often follows an urban-rural hierarchy with changes most often begin in major centers of population then affecting the smaller centers. Hence, the Wave model considers the geographical distance as the most important determinant of linguistic distance.

Trudgill (1974) introduced the Gravity model by further extending this notion. In Trudgill's Gravity model, population size is one of driving forces behind the spread of linguistic innovation (Nerbonne et al., 2005, p.2). Unlike the pure version of the Wave mode, the Gravity model does not postulate that the linguistic innovations simply radiate from one center to another; it rather suggests that the large center is influenced before the smaller one. Hence, it is also called Cascade model (Labov, 2001; Nerbonne et al., 2005). The base of the Gravity model is the frequency of social contact. Accordingly, social connection and adoption of social and cultural innovation is achieved through contact. As social contact increases, the diffusion of innovations increases, and the reverse is true for geographical distance. Trudgill (1974) further suggests that the degree of contact declines, not as a linear function of distance, but rather quadratically. He formulated the relationship between linguistic distance, geographical distance and population size as follow.

$$I_{ij} = s \bullet ((P_i P_j) / (d_{ij})^2)$$

Where,

$I_{ij}$  = the influence of center i on center j

$P_i$  = the population size of center i

$P_j$  = the population size of center j

$d$  = the geographical distance between center i and j

$s$  = index of linguistic similarity or constant needed to allow the change

Hence, according to Trudgill's Gravity model, linguistic distance positively correlates with geographical distance while population size negatively correlates with the linguistic distance. The general prediction of the model is that the larger the center, and the smaller the distance between center and target, the greater the influence on one another. According to Boberg (2000, p.3), two further predictions follow from this. 1. A large center influences another large center before it influences equivalent smaller one. 2. When two centers are not equal in size, the larger will have a greater influence on the smaller one. There have been both positive and negative responses

to Trudgill's Gravity model (see Callary, 1975 for the arguments in favor of the Gravity model, and Bailey et al., 1993; Wikle, 1997; Boberg, 2000 and Horvath & Horvath, 2001 for counter arguments).

Nerbonne et al. (2005) used a dialectometry approach to evaluate Trudgill's Gravity Model. Nerbonne et al. (2005) used Levenshtein distance based on comparisons of spectrograms of sounds. The assumption was that if dialect sites are affected by forces of linguistic gravity, then the pattern observed in synchronic data should reflect the accumulated effect of linguistic gravity. In other words, synchronic differences should reflect historical dynamism. The study further argued that, from a methodological point of view, overall tendencies shown in the diffusion need to be evaluated. In other words, features that have not propagated need to be controlled. As a result, the study used data from earlier times, to make sure that there was completed diffusion and that the commutative effect of social contract was included. Population data before the time of modern mobility was used. Data about population size were taken from the *Geschiedkundige atlas van Nederland; Het koninkrijk der Nederlanden 1815-1931 (Ramaer 1931)*. Walking time cost was taken as the distance between sites. This study reported a significant influence of geographical distance, but little effect of the population size on the linguistic distance. The study did not confirm the gravitation force of population size assumed in Trudgill's Gravity model.

Nerbonne et al. (2005) examined only Low Saxon Dutch dialects. A subsequent study, Heeringa et al. (2011) which included other Dutch dialect areas of the Netherlands including major cities reported the direct relationship between the linguistic similarity and geographical distance. Nerbonne et al. (2005) and Heeringa et al. (2011) are among a few studies that examined the influence of Gravity model using aggregate data. These studies in general suggest two important points. First, a center must be big enough to influence another center. Second, the relationship between linguistic distance and geographical distance is not quadratic as predicted by the Gravity model; rather it is sub-linear (see Nerbonne, 2010).

#### 4.4.2.3 Language Attitude

There are several factors that affect the attitude of speakers of one variety towards other varieties. For instance, Golubović & Sokolić (2013) studied the attitude among Croatian and Serbian speakers. They tested the attitude of the speakers using matched guise test, i.e., the participants listened to the Serbian and Croatian recordings of the fable ‘The North Wind and the Sun’. Then, they rated the characteristics of the person in the recorded speech. The same recordings were also administered in seven other languages which were used as fillers. This study reported that both Croatian and Serbian have negative attitude toward the language of the others. The cause of the negative attitude was attributed mainly to the war the two countries between 1991 and 1995. Croatians consider their Serbian counterparts as aggressive, warmongers and losers just based on the speech sounds they heard while Serbians labeled the Croatians as nationalists, ugly and arrogant. Studies conducted by Abu-Rabia (1996) and Abu-Rabia (1998) on Hebrew-Arabic speakers, in the same manner, indicated that Hebrew speakers consider Arabic speakers as aggressive and violent, and they view Arabic language as a useless language. Arabic speakers have also showed extreme lack of motivation to learn Hebrew.

According to Pavlenko (2006), though there may not be armed conflict and long history of war, political hegemony and nationalism can lead to a negative attitude towards the speakers of the dominant language and towards the language itself. According to this study, for instance, after the disintegration of Soviet Union, many member countries such as Lithuania, Latvia and Estonia re-established their own language and discouraged the use of Russian as a lingua-franca of the region. Russian began to be considered as language of totalitarianism and medium of communist ideology. The same situation can be observed in Ethiopia in which Amharic is, for example, considered as a language of invaders, dictators and medium of feudalism by some members of certain ethnic group largely among the Oromo speakers. As a result, some speakers of Oromo are not interested to learn and to speak Amharic (Wolyie, 2008). In my own study, Feleke (2017), on the attitude of native Amharic speakers towards Ethiopian and Eritrean Tigrigna varieties, the same situation was

observed; Amharic speakers showed a negative attitude towards Eritrean Tigrigna variety probably because of the hostile political atmosphere between the two countries.

However, reports about the effect of attitude on mutual intelligibility have not been consistent. For instance, Van Bezooijen & Gooskens (2007) examined the effect of the attitude of native Dutch speakers towards Frisian and Afrikaans. They employed a questionnaire to investigate the attitude of the speakers towards Frisian and Afrikaans, towards places where these languages are spoken and the speakers of the two varieties. The participants wrote their responses on six-point scales. The study did not report a significant effect of attitude on the Dutch speakers' intelligibility scores. In the same manner, Schüppert & Gooskens (2011) investigated the effect of language attitude on the intelligibility of Danish and Swedish. The aim of the study was examining the association between language attitude and the asymmetry of intelligibility between Danish and Swedish. The participants' task was matching the auditory cognate words they heard with the pictures provided on the computer screen. After the picture labeling task, the participants were asked whether the languages they heard during the experiment were less nice (1), nice (2) or nicer than their native language (3) (see Schüppert & Gooskens, 2011, p.123-127 for the details of the procedures). In this study, no significant correlation was found between language attitude and the degree of intelligibility between the two languages.

Similarly, Schüppert et al. (2015) investigated the relationship between language attitude and mutual intelligibility in Danish and Swedish children and adolescents. They employed a matched guise technique which was also used by Golubović & Sokolić (2013). This study reported a very low positive correlation between language attitude and intelligibility (see Schüppert et al., 2015, p.381-382 for details of the attitude tests). In general, it seems that the relationship between linguistic distance and language attitude is not straightforward. There may be several other factors that determine the relationship between the two.

---

# Chapter 5

## Research Methods

This chapter is about the concrete application of some of the methods discussed in Chapter four. It presents the methods which were applied to address the objectives presented in Chapter one. The contents of the chapter are organized as follow; first, the description of research participants and informants is presented (§5.1). This will be followed by the methods and procedures used to measure the structural distances among the selected Ethiosemitic languages (§5.2). Then, methods used to measure the functional and perceptual distance are explained (§5.3). The explanation includes the procedures used to select the participants, the materials used to design the tests, and the steps followed to administer the tests. This will be followed by the presentation of the methods of clustering and cluster validation techniques. Finally, the methods used to determine the determinants of the linguistic distance will be presented (§5.4).

### 5.1 Assistants and Experimental Subjects

In the present study, the term ‘research assistants’ and ‘experimental subjects’ are used with a slight degree of meaning difference. Research assistants are individuals, especially school teachers, who were involved in selecting the informants (students), preparing test materials - including translating test materials, reading the translated texts during the recordings, and identifying cognates in the sample words. ‘experimental subjects’ refer to individuals (particularly students) who took the tests de-

signed to measure the functional distance and the perceptual distance among the language varieties. The procedures used to select the research assistants and the informants are presented as follow.

### 5.1.1 Research Assistants

The research assistants refer to school teachers who were also bachelor degree holders. They were recruited from nine schools in eight districts in the Gurage Zone and one school from the Silt'e Zone. From each school, three teachers who were native speaker of the variety of the area were selected. In other words, the total of thirty teachers were recruited from the ten schools in the nine districts (eight Gurage districts and one district in Silt'e Zone). The teachers were selected in two screening steps. For the initial screening, the call for participation was announced by distributing printed leaflets in the schools. The leaflets contained a few language requirements such as being the native speakers of the local variety, experience in reading and writing the local variety (since Silt'e is a written language) and lifelong residence in the language area. There were many schools in some of the districts. Except for Mesqan and Gura, a school in the administrative town of each district was selected. Regarding Mesqan, the administrative town is Butajira. Since the residents of Butajira are overwhelmingly Amharic speakers, a school outside the Butajira town was selected. Gura is spoken in Chaha district. Hence, Gura speakers who are from around Gura Megnase (suburban area of Edebir - town of Chaha) were selected.

The call was used to make the first contact with the school teachers. On the leaflets, the phone number and the email address of the researcher were also included so that any interested school teacher could easily get in touch with the researcher. The leaflets were posted on the notice boards of the selected secondary schools. Among the teachers who responded to the call, those who satisfied the requirements in the leaflet were selected for the second screening. The second screenings were conducted using semi-structured interviews. The interviews focused on issues such as the teachers' home language situation, exposure to the neighboring varieties, and language condition in earlier workplaces etc. The interview took place in the



schools of the respective teachers. The interview was administered by the researcher. The teachers (research assistants) who were recruited based on these parameters eventually took part in various tasks including test administration, translation of fables to the local varieties, translation of list of words in the fable, oral presentation of test materials during recording and selection of research participants or students. Each research assistant received a modest payment (300 birr) for their services.

### 5.1.2 The Experimental Subjects

The subjects were selected mainly by the research assistants. Thirty (30) students were recruited from each school. Students of all grade levels (from grade 9 - 12) were included to incorporate as many students as possible. Similar to the selection of the research assistants, the students were recruited in a two-step screening processes. First, all the students who are the native speakers of the local variety were requested to register on a registration form prepared for this purpose. The registration was made by the research assistants. Once the native speakers of a local variety were identified, they were admitted to the second screening. Questionnaires were used for the second screening (see Appendix A.1). The questionnaires contained items about the students' first and second language background, family language conditions, demographic information (age, place of birth, grade level and others) and frequency of their contact with the speakers of other neighboring languages/varieties. The questionnaires were prepared in Amharic since all the secondary school students in the study areas were able to read and write Amharic. Amharic is indeed both the language of schooling and language of the workplace in the study areas except in Silt'e Zone where Silt'e is taught in elementary school. The questionnaires were coded for each school and each study area so that they could be easily identified during the test administration. All the items in the questionnaires were close-ended to maximize the accuracy of the responses; additionally, they were matched with the age and the education levels of the students. The questionnaires were administered by the researcher and the research assistants.

Then, based on the information obtained through the questionnaires, 300 participants (30 from each variety) who speak the varieties of interest as the first language

were selected. Besides, based on the data that were obtained from the questionnaire, it was assured that the participants have lived a significant part of their life in the area where their variety is spoken and that their parents are the native speakers of the variety of interest. Whenever the eligible students that fulfill the requirement exceeded 30 for each variety, 15 male and 15 female students were randomly selected. The functional distance and perceptual distance tests were administered at different time, but in the same language sites. As a result, the number of students who participated in the two tests was not exactly the same. The detailed information about the number of participants is presented in Appendix C.7. In total, 285 participants completed the functional distance test, among which 171 were males and the remaining 114 were females. Moreover, 292 participants took part in the perceptual distance test; 171 were males, and the remaining 121 were females.

## 5.2 Determining the Structural Distance

Phonetic distance and lexicostatistics were used to determine the structural distance among the language varieties. The list of words used for the two measures were obtained from the following sources: a fable, ‘The North Wind and the Sun’, words selected for the Semantic Word Categorization tests, and randomly selected additional words mainly from Etymological Dictionary of Gurage Leslau (1979) and from Völlmin (2017). Total of 240 words were used for the phonetic and lexical comparisons (see Appendix B.1).

Studies previously conducted on Germanic languages used only words in the fable. However, translating the fable from English to the Ethiosemitic languages diminished the number of words in the fable. This is because, in Semitic languages, articles, some of the prepositions, possessive/genitive pronouns and accusative pronouns do not appear independently - they occur either as affixes or clitics. Hence, words collected for the Word Categorization test and additional randomly selected words, from the sources mentioned above, were used to address the issue of the sample size.

The first step in the process of computing the phonetic and lexical distance was

translating and phonetically transcribing the list of words. The list of words were prepared before the fieldwork. The list of words were prepared in English without first translating from English to Amharic to suppress the influence of Amharic on the translators' word choice. If they had been translated from Amharic to the local languages, there could be a direct transfer of the Amharic in the translated versions. Amharic equivalents were provided for the assistants only in a few cases where they are were not able to recognize the English words. The three research assistants from each language area were asked to provide orally their mother tongue equivalent for each English word. Whenever the assistants differ on their responses, they were asked to agree based on majority vote. The final responses of the participants were phonetically transcribed by the researcher. In the process of translation, it was observed that there are concepts that have more than one competing representation. For example, a word 'war' can be either [gaz] or [arib] in Chaha, Ezha and Inor. Whenever this happens, the research assistants were told to choose the most frequent one. Then cognates which are shared across all the language varieties were identified by the researcher and research assistants. The shared cognates were determined based on two parameters: similarity of form and of meaning between the corresponding pairs of words. These parameters were employed in a two-step process of cognate identification. First, the researcher identified pairs of words that share a common root based on the form (phonetic) similarity between the corresponding words. The meaning equivalence between each pair of words was confirmed by the research assistants. In almost all the Semitic languages, sequence of consonants form the basic word meaning (root). Hence, the similarity of the consonantal roots was considered as a core parameter. False friends were excluded using the semantic parameter. Cognates that are shared among six of the ten language varieties were considered for phonetic comparisons. In other words, the phonetic distance was computed among cognates that are shared at least by 60% of the languages. No special attention was given for borrowings since the focus of the present study was on areal classification of the languages.

### 5.2.1 Phonetic Distance

The list of cognates identified using the methods described above were used for phonetic distance measure. The phonetic distance was computed between pairs of cognates based on Kessler (1995, p.65) that there is no need of computing the distance among words that do not have any form similarity. To determine the phonetic distance among the ten South Ethiosemitic varieties, first the phonetically transcribed cognates were aligned. Then, the distance among the cognates was computed using Levenshtein algorithm, based on the number of phones which are inserted, deleted or substituted. The distance computation was made using the simplest cost assignment (see §4.2.2.4.1). The simplest cost assignment assigns equal cost for all the operation. This procedure assigns 1 point for each operation. Levenshtein algorithm provides absolute and relative (normalized) string distance. In the present study, normalized phonetic distance was computed to minimize the unnecessary differences due to variations in lengths of the cognates. The phonetic distance among the cognates was computed using GabMap. Table 5.1 shows samples of phonetic distance between Kistane and Chaha based on the word ‘cloud’. In this particular example, the Levenshtein distance is 2. This distance should be divided by the number of segments (six in this example) to obtain the normalized distance. Hence, the relative or normalized phonetic distance between the two cognates is .33 (2/6). This value can be converted into percentage and presented as 33%.

Table 5.1: Phonetic distance, using Levenshtein algorithm

<b>Kistane - Chaha ‘cloud’</b>					
d	a	m	ə	n	a
d	a	b	ə	r	a
		1		1	
Absolute				2	
Relative				0.33	

Table 5.1 shows the phonetic distance just between two cognates. When the phonetic distance among several cognates must be computed, the Levenshtein algorithm

computes the average distance among all pairs of cognates across the languages of interest. Table 5.2 illustrates this operation based on six randomly selected cognates of Gumer and Chaha.

Table 5.2: Example of phonetic distance

No	Cognates	Gumer	Chaha	Longest Segment	Absolute	Relative
1	began	[k'ənəsəm]	[k'ənəsəm]	7	0	0
2	big	[nik <sup>y</sup> ə]	[nik'yə]	5	1	0.2
3	bed	[alga]	[alga]	4	0	0
4	he	[xut]	[hut]	3	1	0.33
5	saw	[aššəm]	[ažəm]	5	2	0.40
6	were	[banəbo]	[banəwo]	5	1	0.16
	<b>Total</b>			<b>30</b>	<b>5</b>	<b>1.09</b>
	<b>Mean</b>			<b>5</b>	<b>0.83</b>	<b>0.18</b>

Table 5.1 shows that the average absolute distance among the six pairs of cognates is 1.0. The average relative Levenshtein distance is computed by dividing the sum of all relative distances to the number of pairs of cognates, which is  $1.29/6 = .22$  in these particular examples. Hence, the relative phonetic distance, between Gumer and Chaha, based on the above sample data is .22. This value can be converted into percentage and presented as 22%.

### 5.2.2 The Lexical Distance

The lexical distances among the selected ten language varieties were determined by computing the percentage of non-cognates in the total lexical items of pairs of language varieties. The lexical distance was computed based on the 240 words mentioned in the preceding section. Once the cognate and non-cognate words in pairs of language varieties were identified, the lexical distance was computed by dividing the number of non-cognate words by the total lexical items in the pair of languages. Table 5.3 illustrates examples of lexical distance between Muher and Silt'e, on the basis of a list of ten randomly selected words. Among the pairs of words, four pairs are cognates and six pairs are non-cognates. Hence, the lexical

distance is 6/10 (number of non-cognates divided by the total lexical item), which is 0.60. In terms of percentage, the lexical distance between the two varieties based on these examples, is 60%. In Table 5.3, similar superscripts show that the pairs are cognates while different superscripts show non-cognate pairs.

Table 5.3: Example of lexical distance

No	words	Muher	Silt'e	Lexical Dist.
1	added	[dəbbərəm <sup>1</sup> ]	[dəbələ <sup>1</sup> ]	0
2	all	[innim <sup>1</sup> ]	[hullimkə <sup>1</sup> ]	0
3	ape	[k'əmələ <sup>1</sup> ]	[k'əmələ <sup>1</sup> ]	0
4	arm	[hitte <sup>1</sup> ]	[kire <sup>2</sup> ]	1
5	back	[ginžə <sup>1</sup> ]	[čin <sup>2</sup> ]	1
6	barley	[əxi <sup>1</sup> ]	[ixil <sup>1</sup> ]	0
7	basil	[bəssobila <sup>1</sup> ]	[bək'ər <sup>2</sup> ]	1
8	bean	[adəngurro <sup>1</sup> ]	[bolok'e <sup>2</sup> ]	1
9	beautiful	[məlkama <sup>1</sup> ]	[bərəedə <sup>2</sup> ]	1
10	bed	[alga <sup>1</sup> ]	[dugmalə <sup>2</sup> ]	1
	<b>Absolute</b>			<b>6</b>
	<b>Relative</b>			<b>.60</b>

## 5.3 Functional and Perceptual Distances

This section presents tests designed to measure the functional and perceptual distances among the speakers of the ten language varieties. Based on the definition of functional distance and mutual intelligibility provided in 1.2.4, the results obtained from the same test was used to determine both the degree of intelligibility among the ten language varieties and the functional distance among the languages.

### 5.3.1 Semantic Word Categorization

Semantic Word Categorization test (Tang & van Heuven, 2009) was employed to measure the degree of intelligibility among the speakers of the ten varieties of the Ethiosemitic languages. This test was chosen since it was easy to administer for

many languages with a minimal impact of the priming effect, which is the limitation of many intelligibility tests, and of the studies previously conducted on the South Ethiosemitic languages.

### 5.3.1.1 Materials

The material selection and preparation procedures were quite similar to that of Tang & van Heuven (2009). The first step in the material preparation was determining the ten semantic categories to be used for the test. The semantic categories include broad concepts such as plants, fruits, animals, furniture etc. The first step towards the selection of the semantic category was identifying the categories themselves. One of the parameters was the frequency of use of the semantic categories among the speakers of all varieties. For instance, some categories such as musical instruments are extremely culture-specific; as a result, they might not be common among all the speakers. The second parameter was the possibility of a semantic category to incorporate as many lexical items/words as possible. This parameter was important since each semantic category must contain at least ten words. First, the researcher selected the categories, and later they were approved by the research assistants. The categories are cloths, body parts, kitchen utilities, fruits, food type, domestic animals, furniture, vegetables, wild animals and cereals.

A similar parameter was used to determine words to be included under each semantic category. In this case, the frequency of each of the words that can potentially be categorized under each semantic category was computed, and the most frequent ten words were selected for each semantic category. It was not possible to compute directly the frequency of the words though. This was because none of the Ethiosemitic varieties under investigation has a representative and systematically organized corpus, ready for frequency count. Additionally, many of them do not have online oral and written documents which could be used as inputs to create one's own corpus. As a result, a corpus, containing about 100,000 written Amharic words, was created using AntConc software (see Anthony, 2005), and this corpus was used to estimate the frequency of each lexical item. In the corpus, texts of different genres (politics, economics, agriculture, culture, sport, science etc.) were

included to make the corpus as representative as possible. Based on this corpus, words that have relatively high frequency were selected for the test. The frequency was computed based on the assumptions that the frequencies of words of the varieties under investigation are similar to the frequencies of corresponding Amharic words. The assumption was held partly because Amharic and the Gurage varieties are genealogically very related Semitic languages.

Using these words, 10 semantic categories each containing 10 different words (see Appendix B.2) were identified. After the identification of the words and the semantic categories, the words were translated from Amharic to the ten varieties by the research assistants. Whenever there was a disagreement among the three research assistants (the translators), they were told to resolve by majority vote, (2/3). After the translation, each translator pronounced the translated words, 100 words for each variety. The pronounced words were recorded on Adobe Audition running on a personal computer. Then, the three translators from each variety were asked to rate each recording on the Likert scale that ranged from 0 (not natural) to 5 (natural). Then, from the three recordings, the one with the highest rating score was selected for the Word Categorization test.

### 5.3.1.2 The Procedure

In the Semantic Word Categorization test, the participants' recognition capability was tested through multiple choice semantic categorization. In the test, the listeners indicated to which of the ten (10) pre-given semantic categories a spoken word belongs. For instance, the participants were provided with 'banana' and asked to categorize under the correct semantic categories which is 'fruits' in this case. The assumption here was that the correct categorization is achieved only if the listeners correctly recognize the target words. As there were ten semantic categories for each word, the probability of categorizing the words correctly by chance was very small (10%). In the process of applying this test, the first step was organizing the audio input in such a way that the listeners do not hear the same word in the same variety several times. In other words, the priming effect due to the repetition of similar input should to be blocked. Similar to Tang & van Heuven (2009), a Latin Square



system was employed for this purpose. Different data files on compact disks (CDs) were created using the following procedures.

As indicated above, in the Word Categorization test, listeners must not hear the same word repeated. A word which is heard twice or more is more easily remembered than a word which is heard only once - the priming effect. In the present study, there were ten semantic categories, each semantic category consisted of ten words, total of 100 (10\*10) words. Based on these words, different CDs were created. In the first CD, the selected list of 100 words were presented in a fixed random order (1-100) in such a way that two subsequent words are never spoken in the same variety. This is a default order. On the second CD (CD2), the words were presented in the same order except that the presentation begins with the variety in which no.100 of CD1 was spoken which was followed by the varieties in which no. 1 then from no.2 to no. 99 were spoken. Due to this shift, every word on CD2 was spoken in different variety than on CD1. The third CD begins with the varieties in which no. 99 was spoken followed by the variety in which no. 100 of CD 2 was spoken and so on. Through this rotation, a total of 10 CDs, each CD containing 100 words in 10 semantic categories were created (see Appendix B.3 for the order of the words).

One CD was administered to one language area. In the present study, the 100 words on a CD were divided into 10 tracks (sections). Each track was administered to a group of three participants. This means that each track was repeated three times. Since each CD contains ten tracks, a total of 30 students took part on each CD. Tang & van Heuven (2009) used 7 seconds as the response time. In the present study, the response time was increased to 10 seconds in order not to put the students under time pressure. Through this rotation: (1) each listener experienced each word only once. (2) Each listener of each language area heard each word in ten different varieties. (3) Every listener heard one tenth (1/10) of the total words. Before the actual testing, there was a practice session. For this session, a separate CD containing 10 Amharic words and ten semantic categories from additional material was prepared. Each participant practiced at least once before beginning the actual task. More than one practice was allowed based on the confidence and the interest of a participant.

For every track of the CDs, there was an answer sheet (see Appendix A.2). Each answer sheet has its own CD and track numbers (for example, CD 1 Track 2) so that each participant received answer sheet with different code number. Tang & van Heuven (2009) provided the list of ten semantic categories on the response sheet. The same method was used in the present study. Each participant was tested individually in a separate session. After listening to the orally presented words, the participants responded by ticking in a square provided beside each semantic category on the response sheet. The test was administered in quiet classrooms in the selected schools. It was administered by the researcher and one of the research assistants. The intelligibility measure was the number of words correctly matched with the given semantic categories. The test procedure is illustrated by figure 5.1 below.

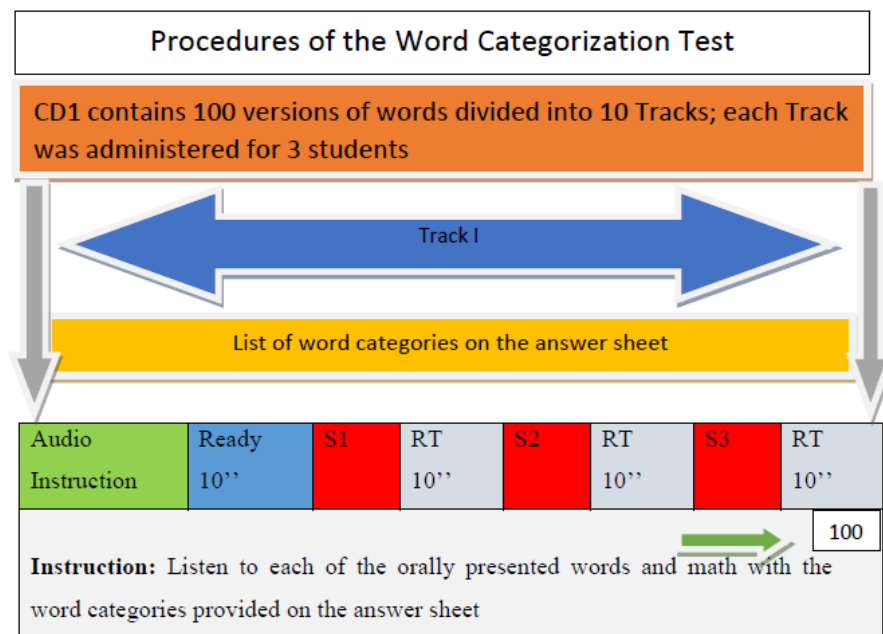


Figure 5.1: Procedures of Word Categorization test; S1, S2, S3 etc stands for stimuli, and RT refers to 'Response Time'

### 5.3.2 Perceptual Distance and Attitude Tests

This section presents procedures which were employed to determine the perceptual distance and the attitude of the speakers towards each others' language variety. The perceptual distance was measured from two perspectives: perceived similarity

and perceived intelligibility. The presentation begins with the materials used for preparing of the tests.

### **5.3.2.1 The Materials**

The fable ‘The North Wind and the Sun’ was used as an input to determine the perceived intelligibility, the perceived similarity and the attitude of the speakers towards each other’s variety. The fable was also used in several previous studies (e.g. Gooskens & Heeringa, 2004 and Tang & van Heuven, 2007). The fable contains simple words which are comprehensible to the speakers of any background. The fable was directly translated from English to the local languages without first translating from English to Amharic to suppress the influence of Amharic on the translators’ word choice. A modified version of Ethiopic writing system was used during the translation from English to the ten local varieties.

Because all the translators were secondary school teachers, they did not face serious challenges in the translation process. During the translation, whenever there was a disagreement among the three translators, the translators were told to resolve the disagreement by majority vote (2/3). After the translation, the translated version of each variety was read aloud by the three research assistants. The reading of each translator was recorded using Adobe Audition running on a personal computer. Then, the three translators listened to each recording and rated the readings on a Likert scale that ranged from 1(not natural) -5 (natural). Finally, among the three readings, the one which received the highest rating was selected for the test. The recording was made in a silent room in each school. The recording process was administered by the researcher.

### **5.3.2.2 The Tests and Test Procedures**

The selected 300 students that were described in 5.1.2 took part in the tests. The perceived intelligibility, perceived similarity and the attitude test items were combined and administered at the same time using the same material. Each test was represented by one item (a question) with its own rating scales. This means that the combined test contains three questions: one for perceived intelligibility; another for

perceived similarity and the remaining one for language attitude. These questions were considered as *test items* in the context of the present study (see Appendix A.3). The three test items were presented simultaneously to minimize the effect of the participants' familiarity with the test material, i.e. the test-takers answered the three questions after listening to each version of the fable.

In order to minimize a response bias that might occur due to fatigue and familiarity with the test materials, the test items were arranged in three different orders; *Order A*: (1) attitude test item, (2) perceived intelligibility test item, (3) perceived similarity test item; *Order B*: (1) perceived intelligibility test item, (2) perceived similarity test item, (3) attitude test item; *Order C* (1) perceived similarity test item, (2) attitude test item, (3) perceived intelligibility test item. Due to these arrangements, each test item appeared in three different orders. Before the test administration, the thirty (30) participants of each language area were randomly divided into three groups, each group containing ten members. Then, the tests were administered in such a way that members of the same group received tests of the same order: one group order A, the second group order B and the third group order C. Administering tests of the same order for members of the same group was important to give the same instruction for all group members. Each group was tested in the same classroom one after another.

During the test, the test-takers listened to the recordings of the fable and responded to the three successive questions. They responded by putting an 'X' on the Likert scale provided to each question. To measure the perceived intelligibility, the participants were instructed to determine to what extent they understand the speaker in the recordings. After listening to each of the recordings, the test-takers indicated their judgment on the Likert scale that ranged from 0 (do not understand at all) to 10 (completely understand). In the same manner, for perceived similarity, the listeners were asked to determine to what extent each of the presented recordings was similar to their own native languages and to put their judgment on the Likert scale that ranged from 0 (not similar) to 10 (completely similar). The perceptual distance was defined as the mean of the participants score on the perceived similarity and perceived intelligibility test items. With regard to the language attitude, the

participants were instructed to determine whether each of the speeches presented in the recordings was beautiful or not, and to provide their responses on the Likert scale that ranged from (1) not beautiful to (10) beautiful. The recordings of the fable in the ten language varieties were presented in different order for the speakers of each variety.

After the presentation of each recording, there was 3 minutes response time, 1 minute for each test item. For the sake of uniformity, the instruction was given in Amharic either by the researcher or by one of the research assistants. Only if there was a serious misunderstanding, further explanation was provided in the participants' native language. The recordings were presented on the personal computer attached to a loudspeaker. After listening to each recording, the listeners answered by putting an 'X' on the scales provided (see Appendix A.3). For each recording, there was a separate answer sheet. In other words, each test-taker received ten pages of response sheets, one page for each recording. This procedure was vital to make sure that the test-takers precisely matched each recording with the respective test items.

### 5.3.3 Clustering and Cluster Validation

After collecting data based on the procedures discussed in the preceding sections, Gabmap was used for the clustering and cluster validation. Gabmap is a web-based dialect classification and visualization software developed by the linguists at the University of Groningen. It provides several statistical alternatives for areal sub-classification of languages (Ward's Method, Complete Link, Group Average and Weighted Average). These alternatives are included in Gabmap based on Prokić & Nerbonne (2008) which evaluated the stability of different classifications techniques. The statistical conception underlying these alternatives is beyond the scope of the present study. In general, based on the sample size and the type of data to be analyzed, one can choose any of the four alternatives. Based on Gooskens & Heeringa (2004), the Weighted Average method was employed to classify the South Ethiosemitic varieties. Clustering is often tricky - a small noise in a data matrix can result in quite different grouping results. Therefore, Gabmap provides two cluster

validation techniques - multidimensional scaling and fuzzy clustering. In the present study, the combinations of these two techniques were employed to make sure that the clustering results are reliable. The two cluster validation techniques are summarized as follow, using sample data from the present study.

Figure 5.2 shows an example of multidimensional scaling. Gabmap provides the result of multidimensional scaling and the corresponding dialect map. The multidimensional scaling plot displays the distance among the languages in an n-dimensional space. In other words, it takes the full site by site distance matrix as an input and creates a representation in the n-dimensional space where distances are approximations of the original linguistic distances (see Leinonen et al., 2016; Nerbonne et al., 1999; Snoek, 2014). The results of the multidimensional scaling can be plotted in a Cartesian coordinate system. On the plot, similar data points are close to each other. The multidimensional scaling output provides options that can be used to refine the classifications based on the nature of distribution of the data. In Figure 5.2, the multidimensional scaling result shows roughly four groups of languages, which are indicated by dots with four different colors. The map on the right, in the same manner, shows four clusters of languages- indicated by the same four different colors. The numbers on the map indicate that two of the languages (Muher and Mesqan) do not precisely fit into the group represented by the green color. The dot on the map represents the average distance of the four languages that precisely fit into the group, with respect to the two languages that do not fit into the group. The number of clusters can be modified by resetting the number of clusters till the most desired classifications are obtained.

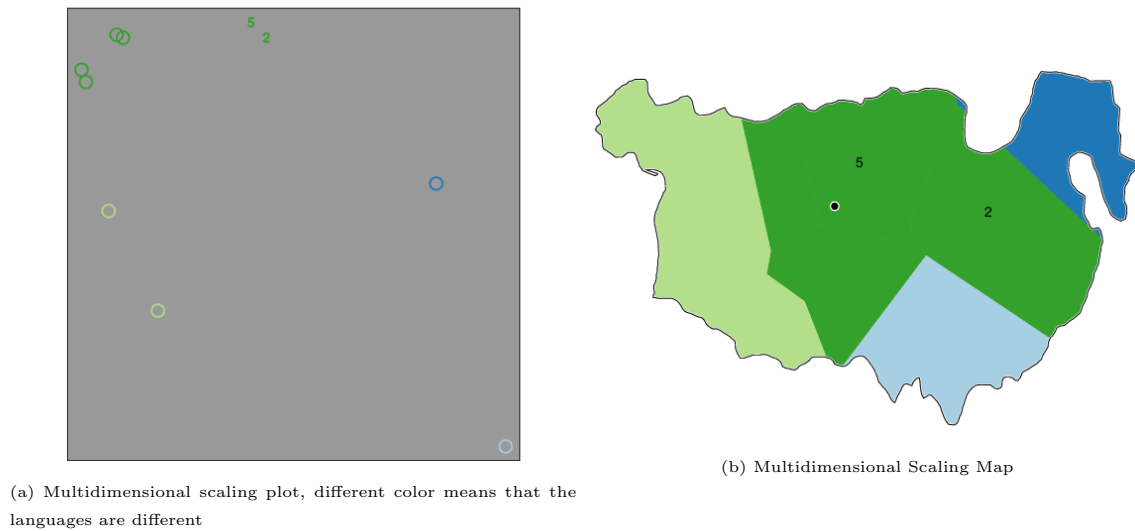


Figure 5.2: Example of multidimensional scaling plot with its corresponding map

The multidimensional scaling results can also be used to inspect the amount of variance explained by each dimension. The first dimension of the multidimensional scaling always explains much of the variance in the data, and additional dimensions are added for the precision of the approximation of the distance, but each additional dimension explains less of the variance than the first one. Figure 5.3 (a), illustrates the plot of multidimensional scaling on a two-dimensional space. The first dimension (solid line) explains 96% of the variance in the data set ( $r = (0.98)^2$ ) while the second dimension (dashed arrow) explains just 2% of the variance ( $(r = .15)^2$ ). The total variance explained by the two dimensions is 98%. Figure 5.3 (b) illustrates the corresponding multidimensional scaling map. On the map, the light color shows area with the highest distance value.

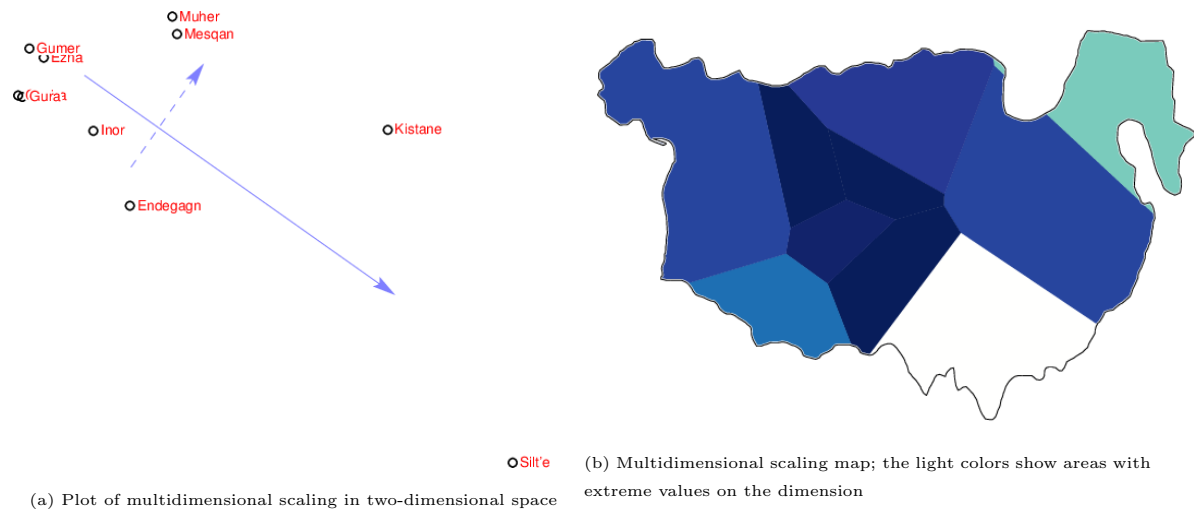


Figure 5.3: Example of multidimensional scaling

In multidimensional scaling, data points with similar values are always shown close to each other<sup>15</sup>. In Figure 5.3 (a), the solid line shows that Gura, Gumer, Ezha and Chaha have the lowest values in the first dimension (the arrow always points from data with low values to those that have high values) while Silt'e and Kistane have the highest values. In the second dimension (dashed arrow), Mesqan and Muher have the highest values and Inor and Endegagn have the lowest values. The plot shows that there are five group of dialects: {Chaha, Gura, Gumer, Ezha}, {Meher, Mesqan}, {Endegagn, Inor} and {Silt'e}.

Figure 5.4 shows an example of fuzzy clustering. In fuzzy clustering, the clustering process is repeated several times after a small amount of noise is added to the data matrix. The numbers on the cluster shows how many times the grouping was confirmed. For instance, '100' on the node that connects Inor and Endegagn means that the grouping was repeated and confirmed 100 times that these two languages form a group. The digits on the scale of the dendrogram indicate the cophenetic distance between these nodes in the group. The map on the right displays the corresponding dialect map. As can be seen from Figure 5.4, the fuzzy clustering shows five groups of languages.

<sup>15</sup>Compare the distance between Muher and Mesqan as opposed to the distance between Kistane and Silt'e .



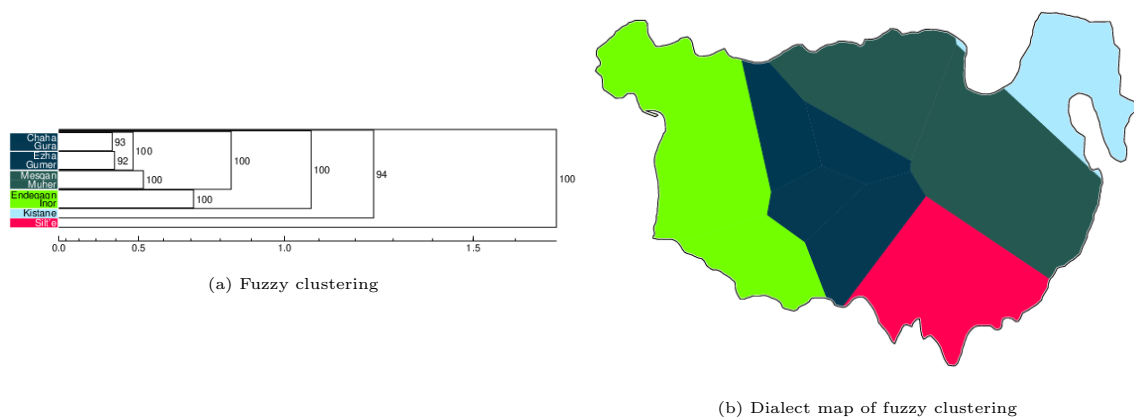


Figure 5.4: Example of fuzzy clustering

After classifying the languages using these techniques (§6.1 - 6.3), the obtained dendrograms were compared to each other in order to derive a comprehensive classification. Two statistical parameters, Local Incoherence and Cronbach's alpha were used as reliability parameters to determine language varieties that must be included in the combined classification. Then, the combined classification was compared to the classifications presented in Chapter three. The cophenetic distance between each nodes was also counted and correlated to each other to examine the similarity between the combined classification and the classification by the historical linguists (§6.4).

## 5.4 Determining Determinants of Linguistic Distance

As indicated in 1.1, identifying some of the linguistic and non-linguistic factors that determine the linguistic distance among the South Ethiosemitic languages is one of the objectives of the present study. In this section, the methods used to detect these determinants are presented. First, the methods of determining the linguistic determinants are presented. This will be followed by the methods of determining the non-linguistic determinants.

### 5.4.1 Determining Linguistic Determinants

The linguistic determinants presented in 4.4.1 were examined using four methods; identifying the characteristic features, examining the horizontal diffusion of features, determining borrowed phonetic and lexical features, and examining the influence of two dominant languages: Amharic and Oromo. As discussed in 4.4.1.1, Gabmap provides statistical options that can be utilized to identify those features which differ little within the group in question and significantly differ between the group. These features were defined as characteristic features or shibboleths. The core concepts of characteristic features are summarized as follow. Given  $g$  = a group of languages to be examined,  $|g|$  = sites of the group of languages to be examined,  $G$  = a large area of interest,  $|G|$  sites in the large area of interest including the sites  $s$  both within and outside  $g$ ,  $d$  = measure of differences between sites with respect to the given feature  $f$ , and  $d^f$  = the set of sites where feature  $f$  is observed, the mean difference with respect to  $f$  within the group is calculated using the following formula.

$$d_f^{\bar{g}} = \frac{2}{|g|^2 - |g|} \sum_{s, s' \in g} d^f(s, s')$$

The mean difference with respect to  $f$  involving elements from outside the group is computed using the following formula.

$$d_f^{\bar{g}'} = \frac{1}{|g|(|G| - |g|)} \sum_{s \in g, s' \notin g} d_f(s, s')$$

The characteristic features were identified using these formulas - they are parts of the Gabmap package. In Gabmap, cognates that have large within-group and outside group differences are ranked on the top of all the cognates used to determine the linguistic distance and those that have low difference are ranked at the bottom of the list of cognates. In the present study, the ten top ranked cognates were used to determine the characteristic phonetic features of each language area illustrated using phonetic distance in 6.1.

In 4.4.1.2.1, it was pointed out that the hierarchical clustering does not distinguish vertical transmission of features from horizontal diffusion of features. Therefore, in the classification of languages using hierarchical clustering, it is not usually

clear whether the similarity among languages is due to the horizontal diffusion or due to the genealogical relationship among the languages. Hence, in the present study, the degree of the influence of contact among the language varieties was determined using the Neighbor-nets algorithm. The algorithm was employed to understand whether the classification among the selected ten South Ethiosemitic languages is due to areal diffusion or not. The Neighbor-net algorithm provides different types of classifications based on the nature of the data matrix. If the input distances show a certain degree of contact, it produces the collection of circular splits - networks. If input distances show a vertical relationship, it will return the corresponding tree. The Neighbor-Net representation was computed based on the lexical distance.

Borrowed phonetic and lexical features were identified using lexicostatistical skewing (Hinnebusch, 1996). The lexicostatistical skewing was used to further examine to what extent the classification of the South Ethiosemitic languages are influenced by the borrowed phonetic and lexical features. As discussed in 4.4.1.2.2, given  $S$  (the similarity index) and given sibling languages, let say,  $L_A$  and  $L_B$ , and another language from another sub-family  $L_D$ , the lexical skewing between  $L_A$  and  $L_B$  with respect to  $L_D$  was defined as follow.

$$S_{AD} - S_{BD}$$

Positive skewing shows borrowing between  $L_A$  and  $L_D$ . Moreover, aggregate skewing of  $L_A$  with respect to  $L_D$  was defined as the average of the skewing between  $L_A$  and each of its siblings with respect to  $L_D$ . As presented in 4.4.1.2.2, the basic assumption of the lexical skewing is that languages which group together tend to have the same distance from another languages in a sister sub-family. Hence, asymmetry between languages in a subgroup relative to languages in another group is considered as an indicator of borrowing. The influence of Amharic and Oromo on the selected South Ethiosemitic languages was determined by examining the percentage of shared phonetic and lexical features between the two languages and the South Ethiosemitic languages. The shared phonetic features among all 240 words was computed using Levenshtein algorithm.

### 5.4.2 Determining Extra-linguistic Determinants

The non-linguistic determinants examined in the present study include population size and geographical distance, contact among the speakers of the languages and the attitude of the speakers. The population size of each language area was obtained from Ethiopian National Census Report 2007. As indicated in 1.1, the Ethiopian National Statistical Agency Census report does not directly provide the number of native Gurage speakers. However, it provides the number of residents in each district. Given that the division of the districts is based on the ethnic background of the residents, the numbers of residents of the ethnically defined districts (e.g., Ezha Wereda, total population 84,905) was taken as the numbers of speakers of the varieties of Ezha. Driving distance (the length of the main road which public transportation (buses) uses) and driving time (time that the public transportation (buses) takes to connect the language areas) between each of the language sites were taken as a measure of geographical distance. Both driving distance and driving time were obtained from Google Map. Phonetic similarity (Appendix C.1) and lexical similarity (Appendix C.2) were taken as measures of the similarity index. The phonetic and lexical similarity indices were computed by subtracting the respective distance measures from 100 ( $s_{\text{index}} = 100 - d$ ). Using these parameters, the interaction between linguistic similarity, population size and geographical distance was examined based on the hypotheses of the Gravity model (Trudgill, 1974). Contact among the speakers was determined using data from the Background questionnaire (see Appendix A.1). The methods used to measure the attitude of the speakers was discussed in 5.3.2.

---

## Chapter 6

# Distance among South Ethiosemitic Languages

In this chapter, results that are obtained from the measure of the three dimensions of distance are presented. The presentation begins with the results of the structural distances (phonetic and lexical distance) (§6.1) which will be followed by the results of the functional distance (§6.2), and then by the results of the perceptual distance among the South Ethiosemitic language varieties (§6.3). The presentation of each result is followed by the presentation of the results of the cluster analysis. The results of cluster analysis are presented together with the results of the multidimensional scaling. As indicated in 5.3.3, Gabmap provides multidimensional scaling in a two-dimensional space in a comparison to cartographic projections.

After the presentations of the results of the cluster analyses, the results of the combined classification of the South Ethiosemitic languages will be reported (§6.4). The combined classification was obtained by comparing the cluster analyses of the three dimensions of distance. As indicated in 5.3.3, Local Incoherence and Cronbachs' Alpha were employed as statistical tools to determine the consistency within the distance matrices and to determine which languages should be grouped together in the combined classification. The combined classification was then compared to the genealogical classifications that were previously proposed by Demeke (2001) and Hetzron (1972). For this comparison, the present study used the cophenetic distance between each pair of language variety (see Gooskens & Heuven, 2018 for the details).

The presentation of the results of the combined classification will be followed by the reports on the relationship among the three dimensions of distance (§6.5). Then, slightly different interpretation of the functional distance will be presented to illustrate the degree of intelligibility among the South Ethiosemitic languages (§6.6). The chapter ends by discussing the major points implied in the results (§6.7).

## 6.1 Structural Distance among the Languages

This section presents the results of the phonetic and lexical distance which were obtained using Levenshtein algorithm and Lexicostatistics respectively.

### 6.1.1 Phonetic Distance

Table 6.1 shows the phonetic distance matrix of the ten South Ethiosemitic language varieties. As discussed in 5.2.1, the phonetic distance was computed using Levenshtein algorithm in Gabmap, based on phones that are inserted, substituted or deleted. The phonetic distance computation was performed on shared cognates. Cognates that are shared at least by 60% of the languages (6/10) were included in the computation. The average value of the Levenshtein distance between each pair of languages was converted to percentages for the sake of ease of presentation. Table 6.1 shows that Chaha has the highest phonetic distance with Silt'e (20%), Endegagn with Silt'e (23%), Ezha with Silt'e (21%), Gumer with Silt'e (22%), Gura with Silt'e (20%), Inor with Silt'e (24%), Kistane with Endegan and Inor (21%), Mesqan with Endegagn (20%), Muher with Silt'e (22%) and Silt'e with Inor (24%). These results show that the phonetic distance between Silt'e and many of the language varieties is high as compared to the distance between the remaining varieties. Table 6.1 also shows that Chaha has the lowest distance with Gura (5%), Endegagn with Inor (14%), Inor with Gumer and Chaha (12%), Gumer with Gura (7%), Gura with Chaha (5%), Mesqan with Chaha (11%), Muher with Kistane and Mesqan (12%), Silt'e with Kistane (18%) and Kistane with Gura (9%). In general, the table shows that Silt'e and Kistane have relatively higher phonetic distances to the other sites, in comparison to other language varieties.

Table 6.1: Phonetic distance among the South Ethiosemitic languages, the lower half of the matrix,  $\alpha = .97$ 

Language <sup>a</sup>	CH	EN	EZ	GM	GU	IN	KS	MS	MU	SI
Chaha	0									
Endegagn	18	0								
Ezha	8	19	0							
Gumer	8	18	8	0						
Gura	5	18	10	7	0					
Inor	12	14	15	14	12	0				
Kistane	18	21	18	19	9	21	0			
Mesqan	11	20	12	13	12	17	13	0		
Muher	14	21	13	15	15	19	12	12	0	
Silt'e	20	23	21	22	20	24	18	19	22	0

<sup>a</sup>ED = Endegagn, IN = Inor, EZ = Ezha, GM = Gurmer, GU = Gura, MS = Mesqan, MU = Muher, SI = Silt'e, KS = Kistane and CH = Chaha. The phonetic distance results were converted to percentage.

Clustering the language varieties based on the phonetic distance matrix in Table 6.1 provides a more comprehensive picture of the phonetic distance among the language varieties. Figure 6.1 (a) shows the multidimensional scaling result of the phonetic distance in a two-dimensional space. The first dimension is indicated by a solid arrow, and the second dimension by a dashed arrow. In multidimensional scaling, language varieties that have similar distance values are placed close to each other in the two-dimensional space while varieties that have different values are placed far from each other (cf: Gura vs. Silt'e). In Figure 6.1 (a), the first dimension (D1) shows that Chaha, Gura, Gumer and Ezha have low phonetic distance values while Silt'e and Kistane have the highest distance values. This dimension explains 52% ( $r = .72$ ) of the variance in the phonetic distance matrix. The second dimension (D2) indicates that Endegagn has the lowest distance value while Mesqan and Muher have the highest values. The phonetic distance values of the remaining varieties are between these two extremes. This dimension explains 38% of the variance ( $r = .62$ ) in the phonetic distance matrix. The two dimensions combined explain 90% of the

variance in the phonetic distance matrix.

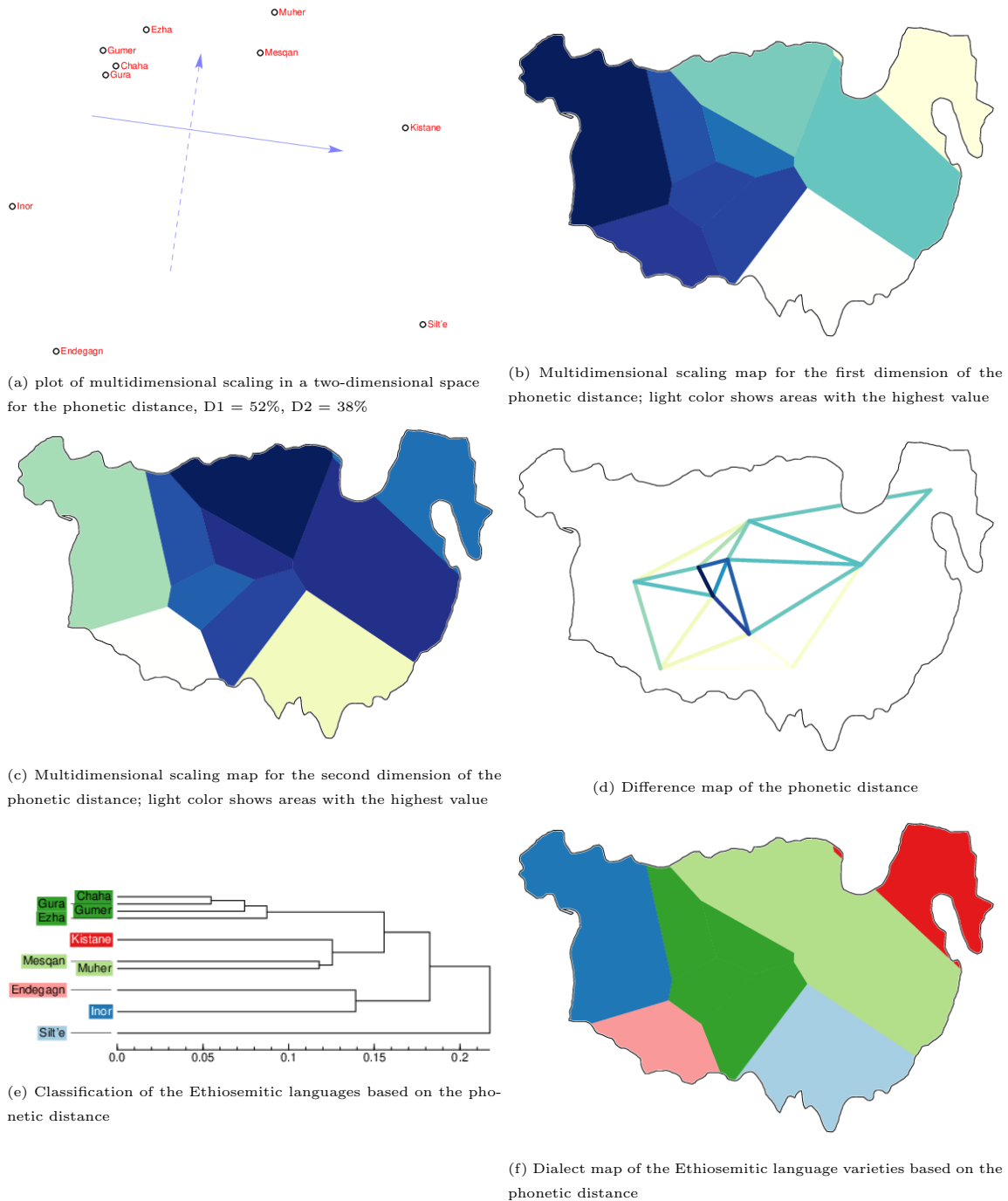


Figure 6.1: Classification of South Ethiosemitic languages based on the phonetic distance

Figure 6.1 (b) shows the map of the first dimension of the multidimensional scaling. On the map, each dialect area is indicated by a different color. The light color represents areas with the highest phonetic distance compared to all other areas, and the dark color shows areas with a low phonetic distance with respect to all



other areas. On the map, it is the Silt'e area that has the highest phonetic distance value. Figure 6.1 (c) shows the map of the second dimension of the multidimensional scaling. The light color shows the area with the highest phonetic distance, which is Endegagn in this case. Figure 6.1 (d) is a difference map which is created by drawing lines corresponding to the aggregate distance between pairs of sites. The darker the color of the line the higher similarity. In this particular map, only neighboring lines are connected by lines. Readers are referred to Nerbonne et al. (2011) and Nerbonne & Siedle (2005) for an alternative map in which lines are drawn across a larger geographic area. The multidimensional scaling results show that, from a phonetic point of view, there are six groups of language varieties - {Chaha, Gura, Gumer, Ezha}, {Mesqan, Muher}, {Endegagn}, {Inor}, {Kistane} and {Silt'e}. Silt'e and Kistane are different from all the other language varieties. This difference was confirmed through several successive classification attempts. Moreover, Endegagn and Inor phonetically behave somehow differently. The results of fuzzy clustering shows almost similar groups of languages (see Appendix C.6).

Figure 6.1 (e) presents the dendrogram which was obtained from the phonetic distance. Varieties which can be clustered together are indicated by the same color. Again, {Gura, Gumer, Ezha, Chaha} form a group. {Mesqan and Muher} are also closely related. {Silt'e} and {Kistane} are separate languages; they do not form groups with any of the language varieties. {Endegagn} and {Inor} are also phonetically different to the extent that they cannot be considered a group. Figure 6.1 (f) shows the geographic extension of six dialect areas. Endegagn, Inor, Silt'e and Kistane form their own separate language area. Hence, both the multidimensional scaling and the clustering results show that, from phonetic point of view, there are six groups of the South Ethiosemitic language varieties.

### 6.1.2 The Lexical Distance

This section presents the results of the lexical distance among the ten South Ethiosemitic language varieties and the corresponding results of the cluster analysis that was performed on the lexical distance matrix. Lexical distance is the percentage of non-cognate words in the total lexical items. Table 6.2 presents the percentage of lexical

distance among the language varieties.

Table 6.2: Lexical distance among South Ethiosemitic languages, lower half of the distance matrix,  $\alpha = .87$

Language <sup>a</sup>	CH	ED	EZ	GM	GU	IN	KS	MS	MU	SI
Chaha	0									
Endegagn	26	0								
Ezha	13	25	0							
Gumer	12	27	13	0						
Gura	11	27	15	12	0					
Inor	22	18	23	20	20	0				
Kistane	39	41	34	33	37	37	0			
Mesqan	24	31	22	21	24	28	30	0		
Muher	21	33	20	20	24	28	28	19	0	
Silt'e	48	50	47	48	49	47	44	46	47	0

<sup>a</sup>ED = Endegagn, IN = Inor, EZ = Ezha, GM = Gurmer, GU = Gura, MS = Mesqan, MU = Muher, SI = Silt'e, KS = Kistane and CH = Chaha; the lexical distance results were converted to percentage.

As can be seen from the table, there is very high lexical differences between Chaha and Silt'e (48%), Endegagn and Silt'e (50%), Inor and Silt'e (47%), Ezha and Silt'e (47%), Gumer and Silt' (48%), Gura and Silt'e (49%), Muher and Silt'e (47%), Mesqan and Silt'e (46%) and Kistane and Silt'e (44%). These results show that the distance between Silt'e and each of the language varieties is higher than the distances among the remaining language varieties. Chaha has the lowest distance with Gura (11%), Endegagn with Inor (18%), Inor with Endegagn (18%), Ezha with Chaha and Gumer (13%), Gumer with Chaha and Gura (12%), Gura with Chaha (11%), Mesqan with Muher (19%), Muher with Mesqan (19%), Silt'e with Kistane (44%) and Kistane with Muher (28%). The results in general show that the lexical distance between Silt'e and the remaining languages is high, compared to the lexical distances among the rest of the language varieties.

Clustering the language varieties based on their lexical distance allows one to visualize these differences better. Figure 6.2 (a) shows the multidimensional scaling result of the lexical distance. The first dimension (D1) indicates that Chaha, Gumer,

Gura and Gumer have low lexical distance values, and Silt'e has the highest lexical distance value. The values of the other varieties are somewhere between these two extremes. This dimension explains the bulk of the variance, 96% ( $r = .98$ ). The second dimension (D2) shows that Endegagn and Inor have the lowest distance values while Muher and Mesqan have the highest distance values. This dimension explains just 2% ( $r = .15$ ) of the variance in the lexical distance matrix. The two dimensions combined explain 98% of the variances in the lexical distance matrix. As can be seen from Figure 6.2 (a), the multidimensional scaling results show five possible groupings of the language varieties: {Gumer, Gura, Ezha, Chaha}, {Muher, Mesqan}, {Inor, Endegagn}, {Kistane} and {Silt'e}. Hence, {Kistane} and {Silt'e} are separate languages. Figure 6.2 (b) shows the map of the first dimension of the multidimensional scaling. Areas with the highest lexical distance are indicated by light color. Again, it is Silt'e that has the highest distance value. Figure 6.2 (c) shows the map of the second dimension of the multidimensional scaling. The lightest area indicates Endegagn area which has the highest distance value. Figure 6.2 (d) is difference map of the lexical distance. Very similar languages are connected by dark lines while dissimilar languages are connected by light lines.

Figure 6.2 (e) presents the dendrogram obtained from the lexical distance matrix. Languages that form a group are represented by the same color. As can be seen from the figure, {Gura, Gumer, Chaha, Ezha} form a group. {Muher, Mesqan} form another group. {Endegagn, Inor} are also very closely related. {Kistane} and {Silt'e} are separate languages. Figure 6.2 (f) presents the dialect map of the language varieties based on their lexical differences. The classifications based on the lexical distance in general show that there are five distinct groups of languages. In the classifications based on both phonetic and lexical differences, Silt'e and Kistane do not form a group with any of the language varieties. It seems that they are different from the remaining Gurage varieties both from the phonetic and the lexical point of view.

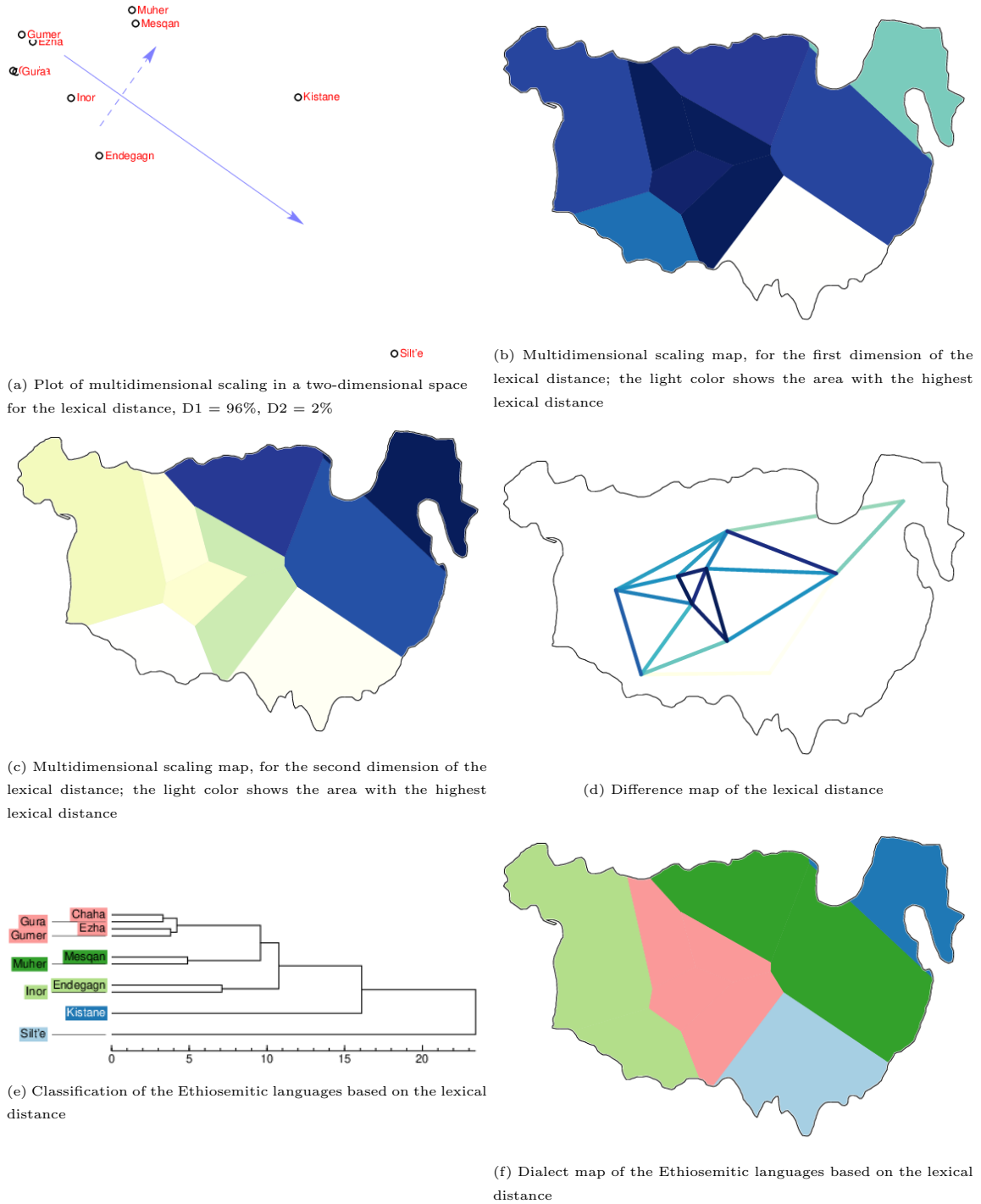


Figure 6.2: Classification of South Ethiosemitic languages based on lexical distance

## 6.2 Functional Distance among the Languages

As presented in 5.3.1, the results of functional distance among the selected South Ethiosemitic languages were obtained from the Word Categorization test. It is

important to recall that the participants' scores on the Word Categorization test were the measures of linguistic similarity not the difference. Hence, the percentage of the respondent's score on the Word Categorization test was subtracted from 100 to obtain the functional distance. For example, if a participant's average score on the Word Categorization test was 60, the functional distance would be  $100 - 60$ , which is 40. Since the functional distance among many of the varieties was asymmetrical (for instance the distance between Endegagn and Inor is 29% while that of Inor and Endegagn is 9%), the upper and the lower halves of the distance matrix were averaged. The average of the two halves of the matrix was computed to minimize the noise in the distance matrix. The participants' scores on their own native language (home language) were excluded to reduce noise in the distance matrix.

Table 6.3: Functional distance among the language varieties,  $\alpha = .63$ 

Language <sup>a</sup>	CH	EN	EZ	GM	GU	IN	KS	MS	MU	SI
Chaha	0									
Endegagn	40	0								
Ezha	20	50	0							
Gumer	17	49	23	0						
Gura	18	53	23	18	0					
Inor	30	19	42	41	45	0				
Kistane	51	52	63	43	47	56	0			
Mesqan	44	58	45	33	50	57	41	0		
Muher	27	53	28	25	28	50	50	48	0	
Silt'e	58	62	56	54	60	73	72	66	65	0

<sup>a</sup>The test languages are abbreviated - CH = Chaha, ED = Endegagn, EZ = Ezha, GM = Gurmer, GU = Gura, IN = Inor, KS = Kistane, MS = Mesqan, MU = Muher and SI = Silt'e ; the functional distance results are converted to percentage.

Table 6.3 presents the functional distance matrix of the South Ethiosemitic language varieties obtained based on the above procedures. The table shows that Chaha has the highest distance with Silt'e (58%), Endegagn with Silt'e (62%), Ezha with Kistane (63%), Gumer with Silt'e (54%), Gura with Silt'e (60%), Inor with Silt'e (73%), Mesqan with Silt'e (66%), Muher with Silt'e (65%), Kistane with Silt'e

(72%) and Silt'e with Inor (73%). Chaha has the lowest distance with Gumer (17%), Endegagn with Inor (19%), Ezha with Chaha (20%), Gumer with Chaha (17%), Gura with Chaha and Gumer (18%), Inor with Endegagn (19%), Kistane with Gumer (43%), Mesqan with Gumer (33%), Muher with Gumer (25%) and Silt'e with Gumer (54%). Like the two structural distances presented in 6.1 and 6.2, the functional distance measure shows that, on average, Silt'e has highest distance from the remaining language varieties.

In order to determine which group of language varieties are different and which are similar, cluster analysis was performed on the functional distance matrix presented in Table 6.3. Figure 6.3 (a) presents the plot of the multidimensional scaling of the functional distance in a two-dimensional space. The first <sup>16</sup>dimension (D1) shows that Silt'e has the highest functional distance value whereas Gumer, Chaha, Ezha and Gura have the lowest functional distance values. The remaining language varieties are between these two extremes. This dimension explains 79% ( $r = .89$ ) of the variance in the distance matrix. The second dimension (broken arrow) shows that Inor and Endegagn have the highest functional distance value, while Muher and Mesqan have the lowest functional distance values. This dimension explains just 14% ( $r = .37$ ) of the variance. The two dimensions, together, explain 93% of the variance in the functional distance matrix. Figure 6.3 (b) shows the map for the first dimension of the multidimensional scaling. The area with the highest functional distance value is indicated by a light color which is, again, the Silt'e area.

Figure 6.3 (c) shows the map of the second dimension of multidimensional scaling for the functional distance. The light color shows areas that have the highest distance values, which are Inor and Endegagn areas. Figure 6.3 (d) is the difference map of the functional distance. Languages that are linguistically very similar are connected by dark lines while the dissimilar ones are connected by light lines. Figure 6.3 (e) shows the dendrograms of the functional distance among the language varieties. Similar languages are indicated by the same color. {Gumer, Gura, Chaha, Ezha,

<sup>16</sup>The participants' score on their native (home languages) were excluded since, in principle, the participants have a perfect knowledge of their own languages.

Muher} form a group, {Mesqan, Kistane} another group; {Endegagn, Inor} also form a group. {Sil'te} is similar with neither of the language varieties. Figure 6.3 (f) illustrates the dialect map of the language varieties based on the functional distance. The figure shows four dialect areas, with Silt'e forming its own distinct dialect area.

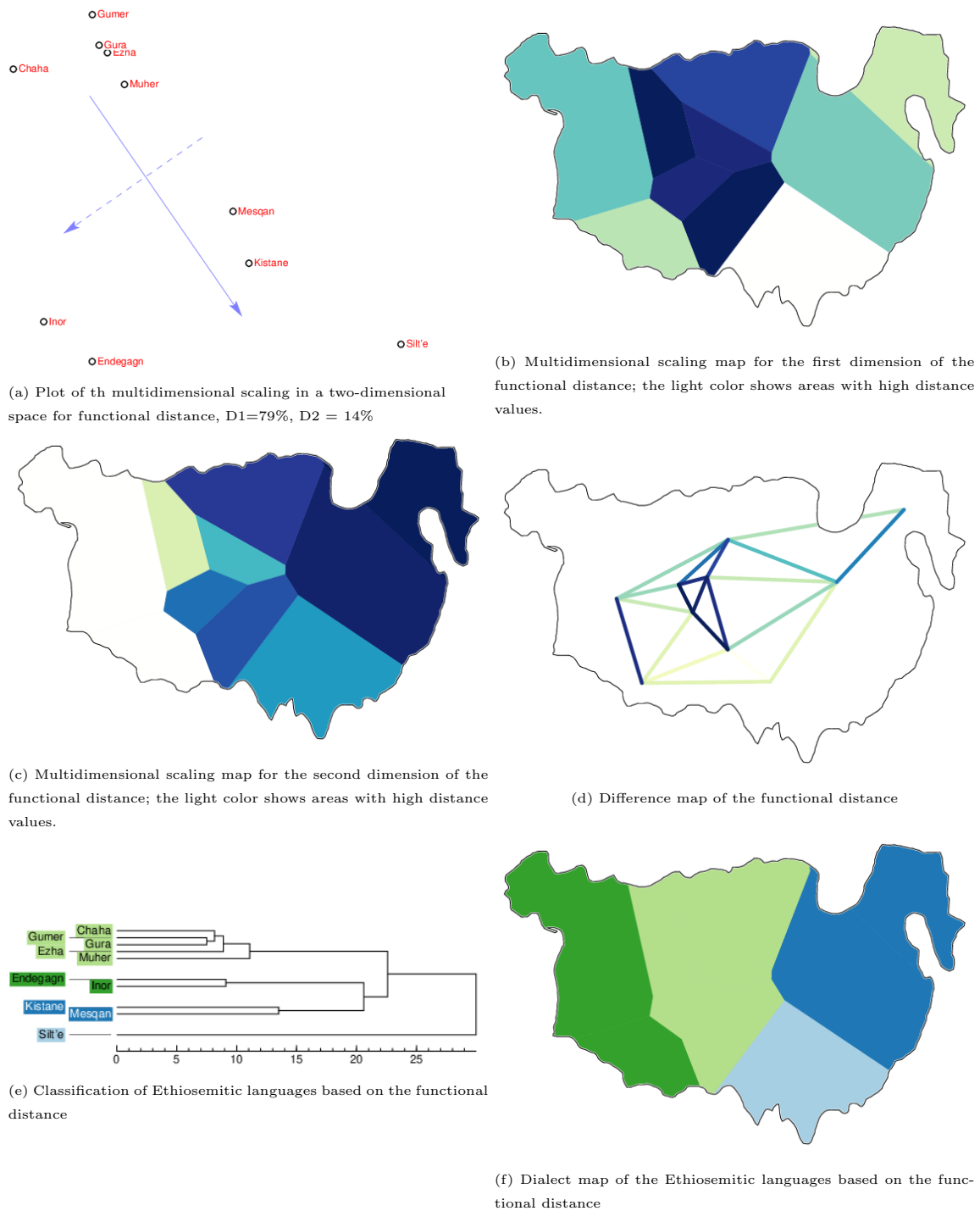


Figure 6.3: Classification of South Ethiosemitic languages based on the functional distance

### 6.3 Perceptual Distance among the Languages

In 5.3.2, it was presented that two perceptual distance measures, i.e., perceived similarity and perceived intelligibility were used to determine the perceptual distance among the language varieties. The mean of the two measures was computed to determine the perceptual distances among the ten varieties. Like the functional distance, the participants' scores on the perceptual similarity and intelligibility test items are measures of the similarity among the language varieties, not the difference. As a result, each similarity score was subtracted from 100 to obtain the perceptual distance. Since the participants' scores were asymmetrical in many instances, the mean of the upper and lower halves of the distance matrix was taken as a measure of perceptual distance. Table 6.4 shows the perceptual distance matrix obtained through these procedures.

Table 6.4: Perceptual distance among the language varieties,  $\alpha = .61$

Language <sup>1</sup>	CH	EN	EZ	GM	GU	IN	KS	MS	MU	SI
Chaha	0									
Endegagn	75	0								
Ezha	31	73	0							
Gumer	14	66	19	0						
Gura	10	57	19	12	0					
Inor	58	19	69	50	44	0				
Kistane	79	55	73	80	80	71	0			
Mesqan	70	64	36	44	41	70	44	0		
Muher	57	74	57	43	46	71	42	32	0	
Silt'e	88	67	79	79	80	86	59	69	71	0

Table 6.4 shows that Chaha has the highest perceptual distance with Silt'e (88%), Endegagn with Silt'e (67%), Ezha with Silt'e (79%), Gumer with Silt'e (79%), Gura with Silt'e (80%), Inor with Silt'e (86%), Kistane with Gumer and Inor (80%), Mesqan with Chaha and Kistane (80%), Muher with Endegagn (74%) and Silt'e with Chaha (88%). The perceptual distance results, like the structural and functional



distance, shows that the distance between Silt'e and many of other language varieties is relatively very high. Chaha has the lowest distance with Gura (10%), Endegagn with Inor (19%), Ezha with Gumer and Gura (19%), Gumer with Gura (12%), Gura with Chaha (10%), Inor with Endegagn (19%), Kistane with Muher (42%), Mesqan with Muher (32%), Muher with Kistane (42%) and Silt'e with Kistane (59%).

Figure 6.4 (a) shows the multidimensional scaling plot of the perceptual distance. As the figure illustrates, in the first <sup>17</sup>dimension (D1), Silt'e has the highest perceptual distance value while Ezha, Gumer, Gura and Chaha have the lowest perceptual distance values. The values of the other varieties are between these two extremes. This dimension explains 76% ( $r = .87$ ) of the variance in the perceptual distance matrix. The second dimension (D2) shows that Inor and Endegagn have the highest values while Mesqan and Muher have the lowest perceptual distance values. This dimension explains 7% ( $r = .27$ ) of the variance in the perceptual distance matrix. Both dimensions, together, explain 83% of the variance in the distance matrix.

Figure 6.4 (b) presents the map of the multidimensional scaling for the first dimension. The light color indicates areas (Silt'e in this case) with the highest perceptual distance values. Figure 6.4 (c) presents multidimensional scaling map for the second dimension of the perceptual distance. The light color indicates areas with the highest linguistic distance; they are Endegagn and Inor areas in this case. Figure 6.4 (d) shows difference map of the perceptual distance. Very similar language areas are connected by dark lines, and dissimilar areas are connected by light lines. The multidimensional scaling results clearly show that there are four groups of language varieties: {Chaha, Gura, Gumer, Ezha}, {Mesqan, Muher}, {Endegagn, Inor} and {Kistane, Silt'e}. From the perceptual point of view, Kistane is very similar to Silt'e. Figure 6.4 (e) and 6.4 (f) show the dendrogram of the perceptual distance, and the dialect map of the ten language varieties respectively. Languages which form a group are indicated by the same color.

<sup>17</sup>In the cluster analysis, participants' score on their own native language, for example Chaha speakers on Chaha was excluded since in principle the native speakers perfectly understand their native languages.

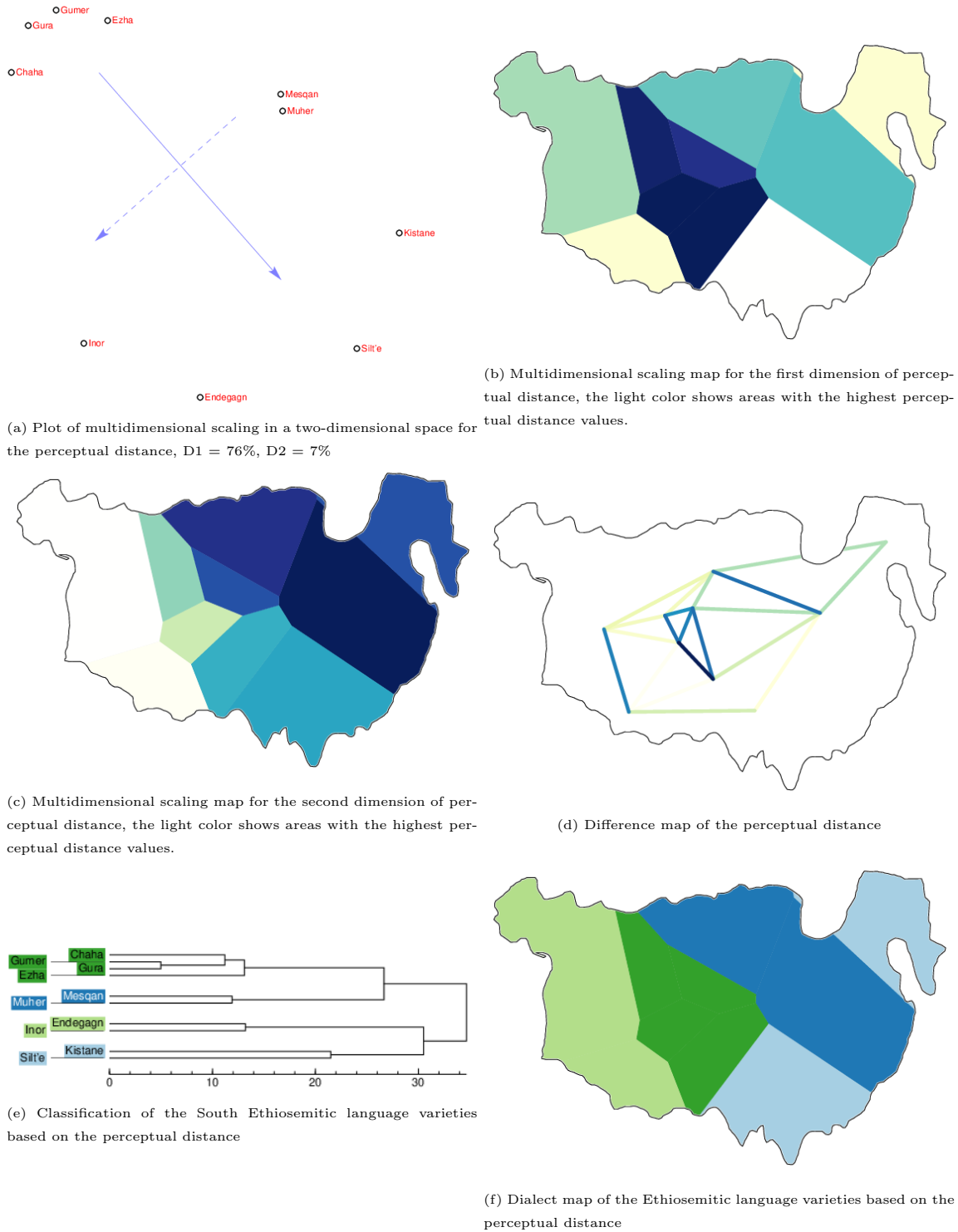


Figure 6.4: Classification of South Ethiosemitic languages based on perceptual distance

Figure 6.4 (e) shows that {Chaha, Gumer, Gura, Ezha} form a group. {Inor, Endegagn} form another group. {Muher, Mesqan} are also very similar. In a manner that is different from the classification based on the structural and functional

distances, Kistane and Silt'e have shown very close similarity in terms of the perceptual distance. Figure 6.4 (f) shows the dialect map of the language varieties based on the perceptual distance among the varieties.

## 6.4 The Combined Classification

As presented in the preceding sections, the classifications that were obtained from the structural, functional and perceptual distance measures are not identical. The classification based on the phonetic distance shows six groups of languages while the classification based on the lexical distance indicates five group of the South Ethiosemitic languages. These classifications are also internally different; for instance, Endegagn and Inor are similar in the classification based on the lexical distance, but they are separate languages in the classification based on the phonetic distance. Hence, this section, aims to combine these classifications and provide a comprehensive classification of the languages. Then the results of the comparison between the combined classification and the classifications by the historical linguists will be presented. Figure 6.5 (a-d) summarizes the dendrograms presented in 6.1 - 6.4. Figure 6.5 (e) presents the combined classification which was derived from the comparisons of all other classifications. The sigma symbols in the combined classification represent unspecified mother nodes.

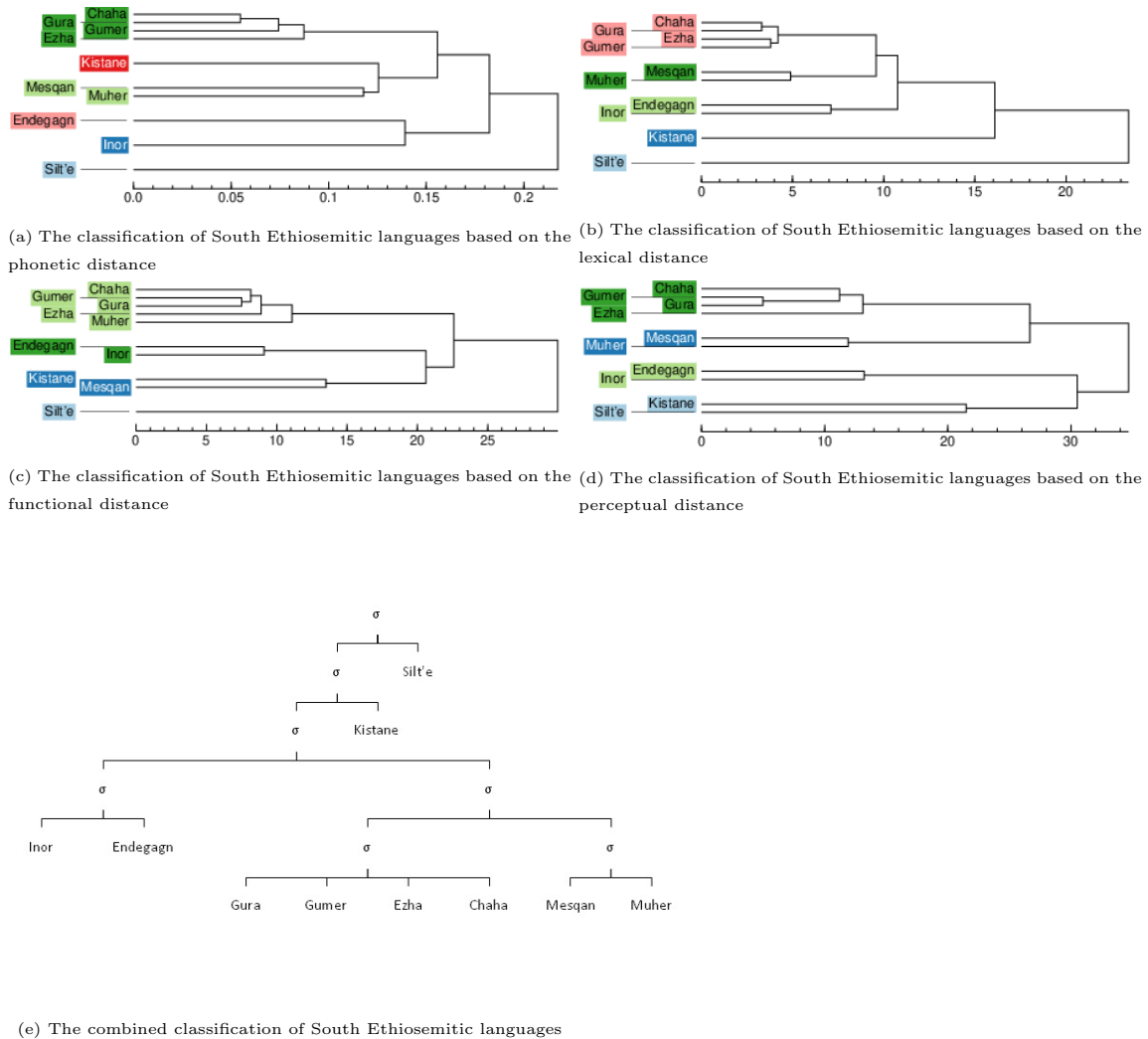


Figure 6.5: Comparisons of clusters obtained from the three dimensions of distance

Given that the linguistic distance was measured from three perspectives (structural, functional and perceptual), the distance matrices were ranked based on their reliability, and the most reliable distance measures were used to combine the classifications presented above. Gabmap provides two measures of reliability of distance matrices: *Local Incoherence* and *Cronbach's alpha*. Local Incoherence is a numerical score of local stress that is assigned to set of differences between items (between languages in the present study). The optimal score is zero while the non-optimal scores can be any positive value. Comparing the value of Local Incoherence for different measurements over the same data gives an idea about which result is more reliable (Nerbonne et al., 2011). Lower values of Local Incoherence mean that the results are better. The idea behind the Local Incoherence is that on average, the locations

that are geographically close should be less different than location that are further apart.

Cronbach's alpha is a coefficient of reliability. It is usually used to measure the internal consistency or reliability of the psychometric test scores. In Gabmap, it is used as the coefficient of reliability of the measurement of differences over the data. High ( $> .70$ ) Cronbach's alpha means that there is high degree of consistency among the measure of distances. Table 6.5 shows the results of Local Incoherence and Cronbach's alpha for each of the distance matrices: phonetic, lexical, functional and perceptual.

Table 6.5: Consistency within the distance matrices

		Local Incoherence	<sup>a</sup> Cronbach's Alpha
Structural	Phonetic	.22	.97
	Lexical	.23	.87
Functional		.29	.63
Perceptual		.32	.61

<sup>a</sup>The high Cronbach's alpha of the phonetic distance could be due to the high sample size. Nonetheless, the relatively higher degree of Cronbach's alpha of the functional distance shows that perceptual distance has extremely low reliability. Also note that the reliability measures for the functional and perceptual distances is based on the mean of the upper and the lower halves of the respective distance matrix.

The phonetic distance has the highest Cronbach's alpha value, and the lowest value of Local Incoherence. This means that it is the most reliable measure compared to all other distance measures. Lexical distance has low Local Incoherence and higher Cronbach's alpha compared to the functional and perceptual distance measures. Compared to the perceptual distance, the functional distance has a high Cronbach's alpha and lower Local Incoherence. Perceptual distance has, by far, the lowest Cronbach's alpha and the highest Local Incoherence which means that it has very low reliability. In general, Table 6.5 shows that the structural distance (both phonetic and lexical measures) are the most reliable distance measures. Functional distance is more reliable than the perceptual distance. Perceptual distance is the least reliable distance measure.

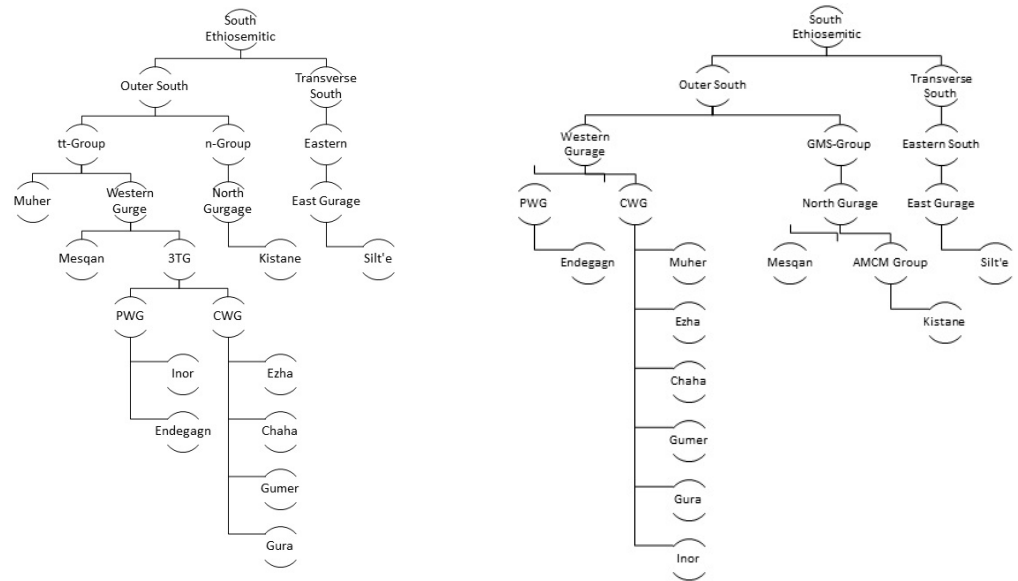
Given these reliability differences, the structural distance was employed as a primary parameter in the process of determining the combined classification, i.e., if a set of the language varieties form a group in both phonetic and lexical classifications that set of languages was automatically considered for the combined classification. However, when languages belong to different groups in the phonetic and in the lexical classification, the functional distance was considered as a second parameter to determine which group is the most plausible one. Perceptual distance was considered as a third parameter when a set of language varieties form different groups in the classifications based on both the structural and functional distances.

In Figure 6.5 (a-d), {Chaha, Gura, Ezha and Gumer} form a group not only in the classification based on the phonetic distance, but also in the classifications based on the lexical distances. Therefore, this group was automatically included in the combined classification without even considering their classification based on the functional and perceptual measures. {Inor} and {Endegagn} are separate languages in the classification based on the phonetic distance, but they are very similar in the classification based on the lexical distance. Hence, the functional distance was used as a second parameter. As Figure 6.5 (c) shows, {Inor and Endegagn} form a group in the classifications based on the functional distance. Based on these requirements, Inor and Endegagn were grouped together. {Mesqan and Muher} form a group in the classifications based on both phonetic and lexical measures. Hence, they automatically qualified for the combined classification. {Silt'e} and {Kistane} are separate languages in the classification based on the phonetic and lexical differences. They are also separate languages in the classification based on the functional distance. Therefore, they were considered as independent languages in the combined classification though they form a group in the classification based on the perceptual distance. This was due to the fact that the perceptual distance has very low reliability. Based on these requirements, the ten South Ethiosemitic language varieties were classified into five groups - {Chaha, Gura, Gumer, Ezha}, {Inor, Endegagn}, {Mesqan, Muher}, {Kistane} and {Silt'e}.

As Figure 6.5 (a-c) shows, the classification of the Central West Gurage languages - Chaha, Gura, Gumer and Ezha is consistent across all the classification parame-

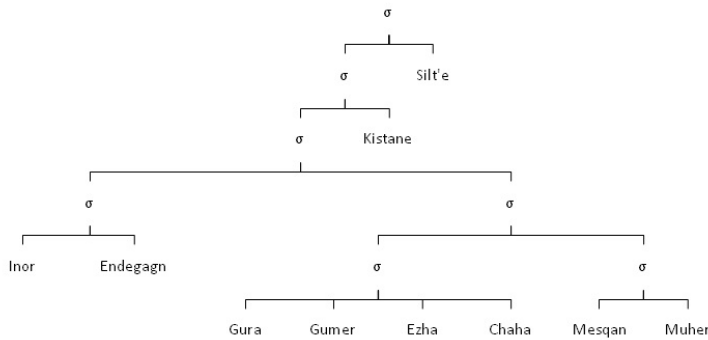
ters. Therefore, these languages were used as a point of reference to determine the relative positions of other groups in the combined classification. {Muher, Mesqan} are close to {Chaha, Gura, Gumer and Ezha} than {Kistane} in the classification based on lexical distances, but in the classification based on the phonetic distance, {Kistane} is rather close to {Chaha, Gura, Gumer and Ezha}. In this case, the functional distance cannot be used as a second parameter since Muher and Mesqan do not form a group in the classification based on the functional distance. Hence, the perceptual distance was used as a third parameter to move {Muher and Mesqan} close to the four Central West Gurage languages. {Inor and Endegagn} are close to the Central West Gurage languages than {Kistane} in lexical, functional and perceptual classifications; therefore, they maintained similar position in the combined classification. Compared to Silt'e, {Kistane} is closer to the Central West Gurage languages based on phonetic, lexical, functional and perceptual parameters. Silt'e is the one that is the most remote from the Central West Gurage languages based on three (lexical, functional and perceptual) of the four classification parameters. The result of this process is the combined classification presented in Figure 6.5 (e).

The remaining point is determining to what extent the combined classification corresponds to the classifications previously proposed by historical linguists. The comparisons of Figure 6.6 (a) - 6.6 (c) shows that the combined classification seems similar to the classification by Hetzron (1972). For example, in both classifications, {Chaha, Ezha, Gura and Gumer} form a group. {Inor and Endegagn} also form a group in both classifications. However, unlike the combined classification, Muher and Mesqan do not form a group in the classification by Hetzron (1972). Moreover, unlike the classification by Demeke (2001), Muher and Inor do not form a group with the Central West Gurage languages - {Chaha, Gura, Gumer and Ezha} in the combined classification.



(a) Classifications of the South Ethiosemitic language according to Hetzron (1972)

(b) Classifications of the South Ethiosemitic languages according to Demeke (2001)



(c) The combined classifications of the South Ethiosemitic language varieties

Figure 6.6: Comparisons between the combined classification and the classifications by historical linguists

Mere impressionistic comparisons of the dendrograms may not precisely depict to what extent these classifications are similar. As a result, the *cophonetic* distance between each node in the classifications was compared to provide statistically sound evidence about the degree of similarity among the classifications. The cophonetic distance between any two terminal nodes in a tree is defined as the number of nodes one has to go up from language A to the lowest common node shared between the member of the pairs and then down to language B (see Gooskens & Heuven,



2018). For example, in Figure 6.6 (c), the cophenetic distance between Muher and Mesqan is two: (1) from Muher one node up to the mother node, (2) from the mother node down to Mesqan. Pearson's correlation coefficient was used to analyze the relationship between the combined classifications presented in 6.4 (e) and the classifications by the historical linguists.

For the sake of simplicity and space, only the ten language varieties under investigation are included in Figure 6.6 among several Ethiosemitic languages previously classified by the historical linguists. Since the distance between the nodes in a family tree is symmetrical (the distance between node A and node B is equal to the distance between node B and node A), the number of pairs of cophenetic distance measures is always  $N*(N-1)/2$ . In the present study, there are 10 language varieties. Therefore, the possible symmetric pairs of languages to which the cophenetic distance has to be computed is  $10*(10-1)/2$ , which is 45. The cophenetic distance between each pair of the South Ethiosemitic languages is presented in Appendix C.5. For the sake of space, only the correlation coefficients between the cophenetic distance of the combined classification and that of the classifications by Demeke (2001) and Hetzron (1972) are presented here. The analyses of the relationship using Pearson's correlation show that the cophenetic distance of the combined classification correlates more strongly to the cophenetic distance of the classification by Hetzron (1972),  $r = .761$  than to by Demeke (2001),  $r = .553$ . The two correlation coefficients are statistically significantly different, Hotelling's t-test,  $t = 6.845$ ,  $p < .001$ .

## 6.5 Correlations among the Three Measures

In this section, correlations among the three dimensions of linguistic distances reported in the preceding sections are presented. Table 6.6 illustrates the correlation coefficients of the two structural distances, the functional distance and the perceptual distance. As can be seen from the table, there is a very strong correlation between the two structural distances - phonetic distance and lexical distance. Furthermore, the correlation between the two structural distances and the perceptual distance is very strong. Compared to other correlation coefficients, the correlation

between the functional distance and the perceptual distance is small. This suggests that the participants' similarity judgment and their actual score on the Word Categorization test may not be exactly the same.

Table 6.6: Correlation coefficients of the three dimensions of distance

		Structural		Functional <sup>a</sup>	Perceptual
		Phonetic	Lexical		
Structural	Phonetic		.874	.804	.853
	Lexical			.849	.777
Functional					.747

<sup>a</sup>The functional and perceptual distance values are obtained by subtracting the values of intelligibility and the perceptual similarity scores from 100 respectively ( $d = 100 - s$ ). The upper and lower halves of the matrices were averaged for the functional and perceptual distances. The participants' scores on their own native languages were excluded.

In general, there are strong correlations among almost all the distance measures compared in Table 6.6. As a result, in Table 6.7, these correlation coefficients are compared to each other to determine if there are statistically significant differences between them. Fisher's z-transformation was employed to compare the correlation coefficients among the three distance measures: structural, functional and perceptual. Table 6.7 illustrates that there are no statistically significant differences among the correlation coefficients of all the distance measures.

Table 6.7: Comparison of the correlation coefficients

Compared Coefficients <sup>a</sup>		z-values	p-values
$r_{PcpD}$	$r_{PD}$ vs. $r_{PcpD}r_{LD}$	1.051	.293
$r_{FD}$	$r_{PD}$ vs. $r_{FD}r_{LD}$	-.654	.513

<sup>a</sup> $PcpD$  = perceptual Distance,  $LD$  = Lexical Distance,  $FD$  = Functional Distance,  $PD$  = Phonetic Distance

## 6.6 Intelligibility among the Gurage Varieties

As indicated in 6.2, both the functional distance and the degree of intelligibility refer to the participants' scores on the Word Categorization test. In other words, the participants' score on the Word Categorization test was used as a tool to determine the degree of distance among the ten South Ethiosemitic language varieties and to measure the degree of intelligibility among the language varieties. In this section, the participants' scores on the Word Categorization test are presented. In 4.2.3, mutual intelligibility was defined as the degree of communication or understanding between the speakers of language A and language B, without having a direct exposure to either of the languages. The assumption in the present study was that the correct categorization of the words into their semantic categories measures the degree of understanding (at the lexical level) of the speakers of the language varieties.

Table 6.8: Mean of the participants' score on the Word Categorization test

Language <sup>a</sup>	CH	EN	EZ	GM	GU	IN	KS	MS	MU	SI
Chaha	81	58	81	85	81	69	50	46	69	42
Endegagn	62	81	48	48	43	71	48	43	57	33
Ezha	80	52	80	76	76	52	36	40	76	40
Gumer	82	54	79	86	82	50	57	68	82	36
Gura	83	52	79	83	86	55	59	59	79	38
Inor	71	91	64	68	55	82	50	45	55	32
Kistane	48	48	39	57	48	39	83	52	35	22
Mesqan	67	42	71	67	42	42	67	85	63	33
Muher	77	38	69	69	65	46	65	42	81	23
Silt'e	43	43	48	57	43	22	35	35	48	87

<sup>a</sup>The test languages are abbreviated - CH = Chaha, ED = Endegagn, EZ = Ezha, GM = Gurmer, GU = Gura, IN = Inor, MS = Mesqan, MU = Muher, SI = Silt'e and KS = Kistane; the intelligibility results are converted to percentage.

To determine the degree of intelligibility among the language varieties, 75% intelligibility threshold was set based on the suggestion of Grimes (1995) and Casad (1987). Hence, 75% and more score in the Semantic Word Categorization test was

considered the confirmation of intelligibility between the test language and the language of the test-takers. 70-74% score was taken as partial intelligibility. Anything less than 70% was considered absence of intelligibility. Table 6.8 show the intelligibility scores of the participants on the Word Categorization test.

As Table 6.8 shows, Chaha speakers understand Ezha (81%), Gumer (85%) and Gura (81%). Endegagn speakers partially understand Inor (71%). Speakers of Ezha understand Chaha (80%) and Gumer (76%). In the same manner, Gumer speakers understand Chaha (82%), Gura (82%), Ezha (79%) and Muher (82%). Gura speakers understand Chaha (83%), Ezha (79%), Gumer (83%) and Muher (79%). Inor speakers partially understand Chaha (71%) and fully understand Endegagn (91%). Besides, Mesqan is partially intelligible to Ezha (72%). Muher speakers understand Chaha (77%). Silt'e and Kistane are not intelligible to any of the language varieties.

As can be seen from Table 6.8, the test-takers did not score 100% on their own native languages although, in principle, it is assumed that the native speakers have a perfect knowledge of their own language. The participants underperformed on their native languages probably due to non-linguistic factors such as fatigue, low quality of the recordings, lack of attentions, noises in the test environment, time pressure and many others. In order to compensate for the influence of these factors, an adjusted mean was computed for the participants' score on the Word Categorization test. It was computed by subtracting the actual mean of the participants' score on their own native language from the hypothetical mean, which is always 100%. Then the mean differences was added to the same participants' score on the non-native languages with the assumption that the factors that affect the participants' score on their native languages equally affect their scores on the non-native languages. For instance, Chaha speakers, in average, scored 81% on their own native languages though they are supposed to score 100%. Therefore, the adjusted mean was computed by subtracting 81% from 100% which is 19%. Then 19% was added to the scores of the Chaha participants on all other language varieties. Table 6.9 presents the adjusted mean scores computed based on the results illustrated in Table 6.8.

Table 6.9: The adjusted mean of the test-takers' score on the Word Categorization test

Language <sup>a</sup>	CH	EN	EZ	GM	GU	IN	KS	MS	MU	SI
Chaha	<b>100</b>	77	100	100	100	88	69	65	88	61
Endegagn	81	<b>100</b>	67	67	62	90	67	62	76	52
Ezha	100	72	<b>100</b>	87	96	72	56	60	96	60
Gumer	96	68	93	<b>100</b>	96	64	71	82	96	50
Gura	97	66	93	97	<b>100</b>	69	73	73	93	52
Inor	89	100	82	86	73	<b>100</b>	68	63	73	50
Kistane	65	65	56	74	65	56	<b>100</b>	69	52	39
Mesqan	82	57	86	82	57	57	82	<b>100</b>	78	48
Muher	96	57	88	88	84	65	84	61	<b>100</b>	42
Silt'e	56	56	61	70	56	35	48	48	61	<b>100</b>

<sup>a</sup>The test languages are abbreviated - CH = Chaha, ED = Endegagn, EZ = Ezha, GM = Gurmer, GU = Gura, IN = Inor, MS = Mesqan, MU = Muher, SI = Silt'e and KS = Kistane; the results are converted to percentage.

Based on the adjusted mean presented in Table 6.9, Chaha speakers understand Endegagn (77%), Ezha (100%), Gumer (100%), Gura (100%), Inor (88%) and Muher (88%). Endegagn speakers freely communicate with Chaha (81%), Inor (90%) and Muher (76%). Speakers of Ezha understand Chaha (100%), Gumer (87%), Gura (96%) and Muher (96%). They also partially understand Endegagn (72%) and Inor (72%). Gumer speakers understand Chaha (96%), Ezha (93%), Gura (96%), Mesqan (82%) and Muher (96%). They also partially understand Kistane (71%). Gura speakers understand Chaha (97%), Ezha (93%), Gumer (97%) and Muher (93%). They also partially understand Kistane (73%) and Mesqan (73%). Inor speakers understand Chaha (89%), Endegagn (100%), Ezha (82%), and Gumer (86%). They also partially understand Gura (73%) and Muher (73%). Besides, Mesqan speakers understand Chaha (82%), Ezha (86%), Gumer (82%), Kistane (82%) and Muher (78%). Muher speakers understand Chaha (96%), Ezha and Gumer (88%) and Gura 84%). Silt'e is not intelligible to any of the language varieties.

As can be seen from Table 6.9, there are asymmetries among the intelligibility

scores. For example, Mesqan speakers understand 82% of Kistane, but Kistane speakers understand 69% of Mesqan. This kind of uni-dimensional measurement is an obvious challenge to language classification since in the classification of languages, it is usually assumed that varieties are classified into non-overlapping sets of dialects. Devising an accurate means of dealing with such non-symmetric measurements will hopefully be the focus of future studies. The present study relies on the suggestion of Tang & van Heuven (2009) that the average of the upper and the lower halves of the intelligibility score is a good predictor of the degree of intelligibility among related languages. Based on this, the mean of the upper and the lower half of the intelligibility matrix was computed to determine the average degree of intelligibility among each pair of the language varieties. The average of the upper and the lower halves of the matrix of the adjusted intelligibility scores is presented in Figure 6.7.

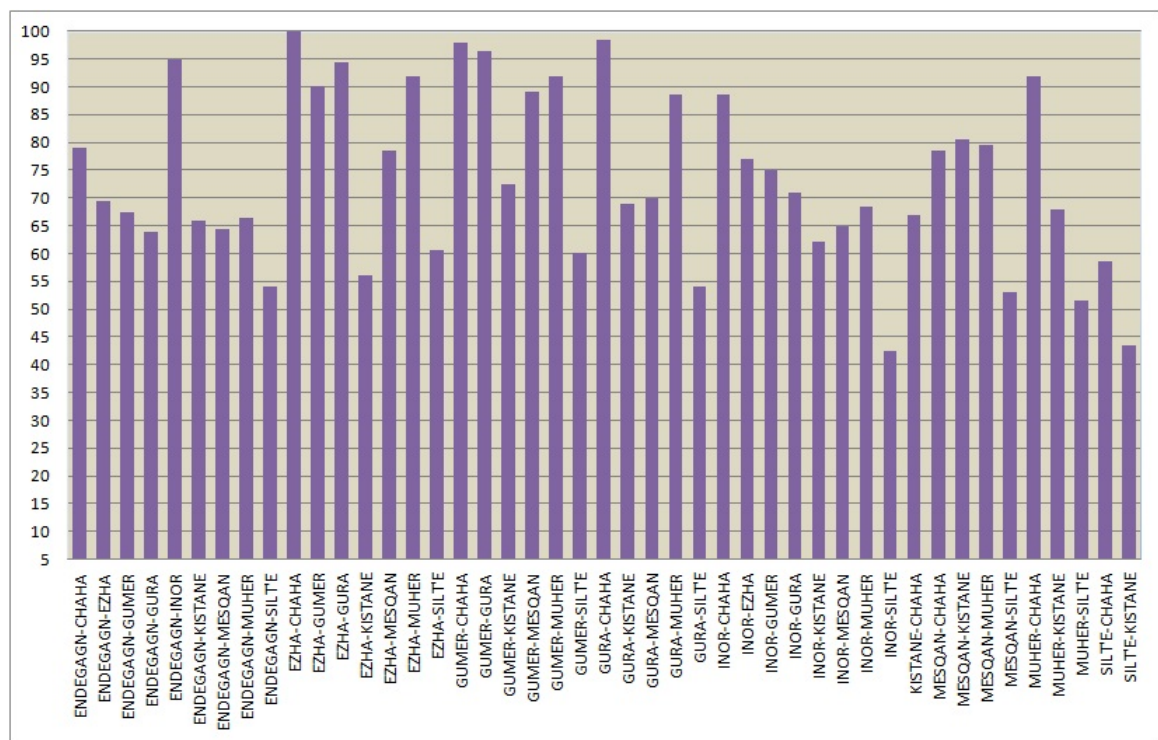


Figure 6.7: The mean of the upper and lower half of the intelligibility matrix for the intelligibility among the South Ethiosemitic languages

The figure shows that, based on 75% intelligibility threshold, Endegagn and Chaha, Endegagn and Inor, Ezha and Chaha, Inor and Gumer, Inor and Gura, Inor and Mesqan and Inor and Muher are mutually intelligible. Moreover, Gumer and

Ezha, Gumer and Gura, Gumer and Mesqan and Gumer and Muher are mutually intelligible. Besides, Gura and Chaha, Gura and Muher are mutually intelligible. Inor is intelligible with Chaha, Ezha and Gumer. Mesqan is intelligible with Chaha, Kistane and Muher. Muher is intelligible with Kistane.

Menuta (2015) argues that the best center of communication is Mesqan, based on the study he conducted on six Gurage varieties - Chaha, Inor, Kistane, Mesqan, Muher and Wolane. This means that many speakers of Gurage varieties understand Mesqan better than the remaining Gurage varieties investigated in the study. The present finding contradicts this report. As can be seen from Figure 6.8, it is Chaha that seems to be the center of communication. Chaha is intelligible to seven of the ten language varieties investigated in the present study. Silt'e was excluded from the Figure since it is not mutually intelligible to any of the language varieties. In Figure 6.8, the two-directional arrow shows that the intelligibility is symmetrical while one-directional arrow shows that the intelligibility is asymmetrical.

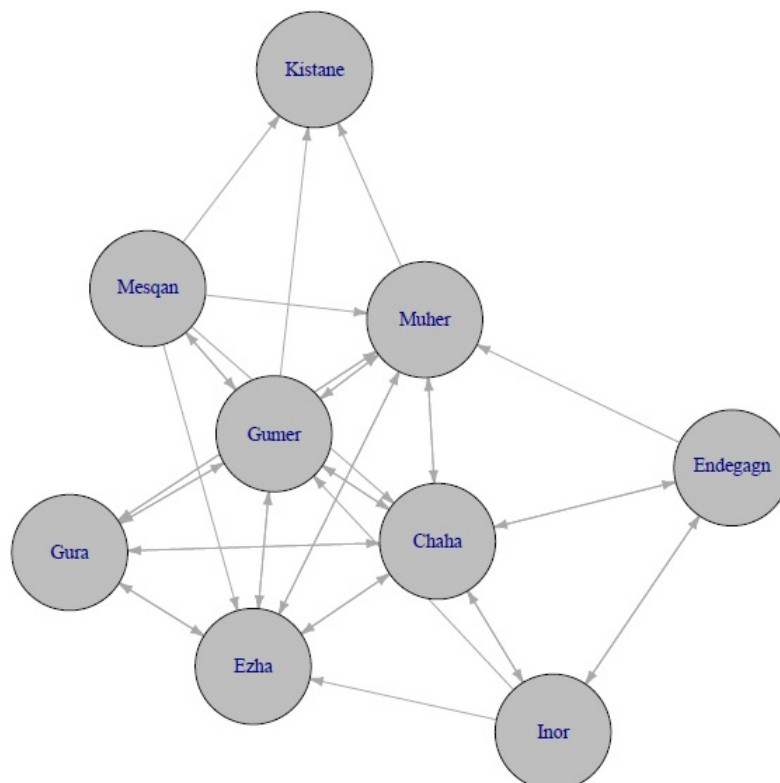


Figure 6.8: Chaha is selected as a center of communication. The two dimensional arrow shows a symmetry of intelligibility while one-directional arrow shows an asymmetry of intelligibility between pairs of languages.

The difference between these two findings can be related to various factors. First, the present study tested the intelligibility at a word level. The author recognizes that testing intelligibility at higher linguistic levels may yield different results. Nonetheless, the present study opted for the inclusion of relatively a large number of languages and examine them from different perspectives rather than focusing just on intelligibility. In this regard, Menuta (2015) included several tests, which is very positive. Nonetheless, there are also concerns about the approaches of Menuta (2015). It appears to me that the priming effect was not properly controlled; since the same test materials were repeated across the speakers of the varieties, it is possible that the intelligibility scores were inflated because of the participants' familiarity with the test materials. Besides, Menuta (2015) tested elderly people while the participants of the present study are secondary school students. It could be the case that elderly people performed on some of non-native languages better than the youngsters mainly because of the lifelong exposure they have had to the non-native language varieties. Sample size could also be another factor. Menuta (2015) tested 12 participants from each site. The present study tested 30 participants from each site. Carefully selecting a small sample size could probably lead to exceptional performance because of the exceptional linguistic abilities of the participants. Moreover, during test administration, Menuta (2015) asked the participants to provide written answers. It is not clear how the participants managed to provide written answers since none of the Gurage varieties (except Silt'e) has a writing system.

## 6.7 Discussions

The classifications of the Gurage varieties based on the results obtained from the structural, functional and perceptual distance measures show that Chaha, Ezha, Gumer and Gura are very closely related languages. Mesqan and Muher are also very similar. In the same manner, Inor and Endegagn are closely related. Kistane has some affinity with Mesqan and Muher, but quite different from many of the Central West Gurage languages. Silt'e is different from all other language varieties. The sub-classifications proposed here are very similar to classifications previously



proposed by Hetzron (1972), but somehow differ from the proposal of, for example, Demeke (2001). For instance, Demeke (2001), based on morphological parameters, classified Mesqan under North Gurage together with Kistane. He also classified Muher with Central West Gurage languages (Chaha, Ezha, Gumer, and Gura). Moreover, although from structural and functional point of view, Kistane and Silt'e are quite different languages, the speakers of the two language varieties think that their languages are similar to each other. This mismatch could be due to the influence of various extra-linguistic factors. Some of these factors will be elaborated in Chapter seven.

With regard to the intelligibility among the Gurage varieties, the results obtained from the Word Categorization test show that Chaha, Gura, Gumer and Ezha are mutually intelligible. The intelligibility between Chaha and Gumer was previously reported in Ahland (2003). Mesqan and Muher are also mutually intelligible with the four Central West Gurage languages. The intelligibility between Chaha and Muher was previously reported in Menuta (2015). Inor is intelligible to Endegagn speakers and with the four Central West Gurage language varieties. The same result was reported in Ahland (2003). Muher and Mesqan speakers understand Kistane somehow. Silt'e is not intelligible to any of the Gurage varieties investigated in the present study. Menuta (2015) argues that Mesqan is the most understood language in the Gurage area. As presented in 6.6, this claim could not be substantiated by the present study. This difference could be due to various reasons as discussed in 6.6. This implies that the degree of intelligibility among these language varieties is an area that requires further rigorous investigations. The reported intelligibility scores are generally asymmetrical. As noticed by Bleses et al. (2008), Gooskens (2018), Gooskens et al. (2010) and Gooskens (2007), this asymmetry could be due to various linguistic and non-linguistic factors. Some languages can be incomprehensible to the speakers of other languages because of their complex linguistic features. For instance, various reduction phenomena such as schwa reduction may lead to the asymmetry of intelligibility (see Bleses et al., 2008). Gooskens (2018) has also discussed various non-linguistic variables such as contact and experience, orthography, gesture and language attitude. In general, as there are several contributing

factors, it is only through the detailed investigation of these factors that the nature of intelligibility among any related languages is fully understood.

As presented in 6.5, the comparisons among the measures of the three dimensions of distance show that the two structural distances (phonetic and lexical) strongly correlate with each other. This implies that the two structural measures can be used interchangeably to determine the linguistic distance among related languages. The present study also reported very strong correlation between the structural distance and the functional distance though different materials were used to measure the two dimensions of distance. This suggests a high degree of possibility of substitutability between the two dimensions of measuring linguistics distance. The strong correlation between the structural distance and the functional distance indicates that the participants' scores on the Word categorization test has a strong connection with the properties of the structure of the language varieties. Given that there was no significant difference between the correlation coefficient of the phonetic distance and intelligibility scores and that of the lexical distance and intelligibility scores, the participants' scores on the Word Categorization test is probably due to both the lexical and phonetic similarity among the language varieties. This may further imply that both the phonetic and the lexical similarities among the languages equally contribute to the speakers' understanding of the language varieties.

The strong correlation between the two structural distances and the functional distance which is reported in the present study is slightly different from previous studies which reported a stronger correlation between the lexical distance and functional distance as compared to the correlation between the phonetic distance and functional distance (e.g., Tang & van Heuven, 2009; Tang & Van Heuven, 2015), and from the studies which reported a stronger correlation between the phonetic distance and functional distance but not between the lexical distance and functional distance (e.g., Gooskens & Van Bezooijen, 2006); hence, the results of previous studies contrasting sharply. Maybe, there are many factors such as similarity of phoneme inventories and frequency of the lexical items which contribute to the relationship between the functional distance and the structural distances. The relationship between these two dimensions is also probably a language specific one. For instance,

in some languages, a slight phonetic difference may lead to misunderstanding, in others it may not.

What is more, the strong correlation between the perceptual distance and the remaining distance measures shows that the perception of the non-linguists can be used to determine the distance among related languages (the assumption here is that the small Cronbach's alpha of the perceptual distance is due to a small sample size). Similar results were also previously reported by several studies (see §4.3). This is good news particularly for less studied languages that do not have dictionaries or detailed descriptions of their linguistic features. However, the high degree of inconsistency in the perceptual distance matrix (low Chronbach's alpha and high Local Incoherence) hints that there is a risk of using a mere perceptual distance to measure the linguistic distance among related languages. The present study shows that the perceptual perspective of measuring linguistic distance is more subjectivity-prone than other means of measuring linguistic distance. As will also be presented in Chapter seven, the perceptual distance is highly sensitive to the attitude of the speakers. As noticed by Golubović & Sokolić (2013), Abu-Rabia (1996), Abu-Rabia (1998) and Pavlenko (2006), the impact of language attitude is more pronounced in situations where there are political, social and cultural antagonisms.

Furthermore, the close similarity between the classifications based on the three dimensions of distance and the genealogical classifications previously provided by the historical linguists implies that, in addition to the structural distances, functional and perceptual distances can be used to classify related languages. In the present study, a very close similarity between the areal classifications and the genealogical classifications has been noticed. This result is consistent with previous reports by Gooskens & Heuven (2018) and Tang & van Heuven (2009). Nevertheless, the similarity between the areal and genealogical classification raises several theoretical issues. These issues will be discussed in Chapter eight. In general, the correlations among the three dimensions of distance which are reported in the present study are consistent with the studies previously conducted on Scandinavian languages (e.g., Gooskens & Heeringa, 2004; Gooskens, 2005; Gooskens, 2007; Gooskens, 2018 and on Chinese dialects (e.g., Tang & van Heuven, 2007, Tang & van Heuven, 2009; Tang

& Van Heuven, 2015). These studies, in general, indicate that the distances among related languages can be measured from different perspectives as far as various intervening extra-linguistic variables are properly controlled.

The present study partially supports the claim that the non-linguists' linguistic consciousness can be used as a means of measuring distances among related languages, contrary to the long-standing debate about the validity of the perception-based approach (see Goeman, 1999 for the debate). The present study also recommends examining the sociopolitical and cultural settings of the speech communities before employing the perceptual distance as a means of measuring linguistic distance. It could be that the perceptual perspective of measuring linguistic distance is more effective in the situations where there are less political, cultural and social tensions. This issue will be discussed in Chapter seven.

Based on the similarity and comprehensibility results presented throughout this chapter, one can conclude that some Gurage varieties can function as variants of common languages. This conclusion is entirely based on the areal similarity among the languages, and it is not to imply the genealogical relationship among the languages. Though providing a clear-cut boundary between a dialect and a language is always difficult due to various linguistic and non-linguistic factors, especially the results of the cluster analyses and the Word Categorization test scores suggest that Kistane is somehow different from many of the Gurage varieties. Silt'e is completely different language from all the varieties. The remaining languages are possibly dialects of the same language. Determining whether these varieties are dialects or independent languages may constitute a significant contribution to the attempts that have been made to standardize the varieties. The results of the study indicate that Kistane and Silt'e need to be treated as separate languages in the standardization process. The remaining Gurage varieties can be considered as dialects, and the same teaching materials can be used to employ these languages in schools, media and in different administrative contexts.

---

## Chapter 7

# Determinants of Linguistic Distances

In Chapter six, the classifications of the ten South Ethiosemitic languages were presented. In the same chapter, it was also illustrated that the classifications of the languages which were obtained from different distance measures are not identical. In this chapter, some of the factors contributing to those classification differences are examined. For the sake of space, this chapter focuses only on some of the factors, particularly on the major determinants discussed in 4.4. In 4.4, two types of determinants were explained - linguistic and non-linguistic determinants. The two major linguistic determinants are characteristic features (shibboleths) and contact-induced phonetic and lexical features. Shibboleths were defined as phonetic features which characterize a group of languages. They were identified using Gabmap. Contact-induced phonetic and lexical features were identified by re-examining the pattern of the classifications using the Neighbor-net algorithm, by computing the percentage of borrowed phonetic and lexical features and by examining the influence of Amharic and Oromo on the South Ethiosemitic varieties. As presented in 5.4.1, lexicostatistical skewing was employed to estimate the percentage of borrowed phonetic and lexical features (not to determine the loaned features). As to the non-linguistic determinants, the present study focuses on three major factors (see §5.3): the degree of contact among the speakers, population size and geographical distance, and the attitude of the speakers towards each others' language. The presentation begins with the linguistic determinants; then the non-linguistic determinants will follow.

## 7.1 Linguistic Determinants

The classifications presented in chapter Six were obtained from structural, functional and perceptual measures; however, the characteristic features cannot be determined for all these measures since, for example, functional and perceptual distances were expressed in terms of numerical values. Therefore, the characteristic features were determined only for the classifications based on the phonetic distance since Gabmap provides support to detect such phonetic features. The influence of contact-induced features was illustrated using a Neighbor-net representation. Regarding borrowing, the amount of skewed features was determined for both phonetic and lexical similarity indices.

### 7.1.1 Characteristic Features

In 6.1.1, six groups (clusters) of South Ethiosemitic language varieties were identified on the bases of the phonetic differences among the languages; {Silt'e}, {Kistane}, {Endegagn}, and {Inor} were considered as separate languages. {Ezha, Gumer, Gura, Chaha} form a group while {Mesqan, Muher} form another cluster. This section presents phonetic features that are responsible for the formation of the later two clusters. Since the first four are independent languages, in principle, they do not have an adequate number of special characteristic features that make them similar to other languages. As presented in 3.2, {Ezha, Gumer, Gura, Chaha } were classified under 'Central West Gurage' by Demeke (2001) and Hetzron (1972). For the sake of convenience, this denomination is adopted in this section too. {Mesqan and Muher} fall under the 'tt-group' in the classification by Hetzron (1972). As there is a debate about this classification (see Demeke, 2001), in the present study, the two languages are simply named *West Gurage languages*.

It was indicated in 4.4.1 that Gabmap provides statistical options that can be used to identify those features which differ little within the group in question and significantly differ outside that group. The core concepts of characteristic features discussed in 4.4.1.1 are summarized as follow; the mean difference with respect to  $f$  within the group is calculated using the following formula.

$$d_f^{\bar{g}} = \frac{2}{|g|^2 - |g|} \sum_{s, s' \in G} d^f(s, s')$$

The mean difference with respect to  $f$  involving elements from outside the group is computed using the following formula.

$$d_f^{\bar{g}'} = \frac{1}{|g|(|G| - |g|)} \sum_{s \in G, s' \notin G} d_f(s, s')$$

Using these notions, Gabmap identifies the characteristic features with relatively large differences between  $d_f^{\bar{g}'}$  and  $d_f^{\bar{g}}$  (see §4.4.1.1). In Gabmap, linguistic features (cognates in the present study) that have large differences are ranked on the top of the list and those that have low difference are ranked at the bottom of the list. The top ranked ones are words that have large outside-group differences, but small within-group differences. These words also contain important phonetic features that characterize each of the two Gurage areas: Central West Gurage and West Gurage languages.

Table 7.1 shows the characteristic phonetic features of top ranked ten words of the Central West Gurage language varieties. The pronunciations of the words in the Central West Gurage area are compared to the pronunciations of the same words elsewhere outside the Central West Gurage. As can be seen from Table 7.1, the distinctive phonetic features of the area are indicated using bold fonts. The table shows that the main phonetic features that make the Central West Gurage language varieties distinct from other Gurage varieties include labialized velar fricative [x<sup>w</sup>], velar ejective [k'], alveolar stop [t], alveolar nasal [n], palatal ejective [c'], labialized nasal bilabial [m<sup>w</sup>], bilabial plosive [b] and alveolar trill [r].

Table 7.1: The pronunciation of ten top ranked words in Central West Gurage varieties (Chaha, Gura, Gumer and Ezha)

Rank	Words	In the Cluster	Not in the Cluster	Characteristic Features
1	now	[ə <b>x</b> <sup>w</sup> a]	[a <b>kk</b> u(ə)]/[a <b>kk</b> ə]/[wa <b>ka</b> ]	[x <sup>w</sup> ]
2	hundred	[bək' <b>i</b> r]	[bəʔ <b>i</b> l]/[bəʔ <b>ə</b> r]	[k']
3	she	[x <sup>y</sup> <b>i</b> t(a)]	[ki <b>y</b> a]/[ki <b>d</b> a]	[t]
4	had	[b <sup>w</sup> n(n) <b>ə</b> n]	[b <sup>w</sup> annən(n) <b>i</b> t]/[ban <b>ə</b> ndə]	[n]
5	frog	[k' <sup>w</sup> ən <b>c'</b> ə]	[k' <sup>w</sup> ən <b>č</b> ə]/[k' <sup>w</sup> ə <b>č</b> ə]	[c']
6	who	[ <b>m</b> <sup>w</sup> an]	[ <b>m</b> a(n)]	[m <sup>w</sup> ]
7	cloud	[da <b>b</b> əra]	[da <b>m</b> əna]/[da <b>w</b> əna]	[b, r]
8	he	[ <b>h</b> ut]/[x <b>u</b> t]	[k <sup>w</sup> <b>a</b> ]/[hu <b>d</b> ə(a)]	[t]
9	were	[ban(n) <b>ə</b> bo]	[ban(n) <b>o</b> wə <b>tt</b> ə]/[ban <b>i</b> m <b>u</b> tt]	[b]
10	our	[y <b>i</b> na]	[in <b>y</b> ]/[y <b>ə</b> ñ( <b>ñ</b> )a]	[n]

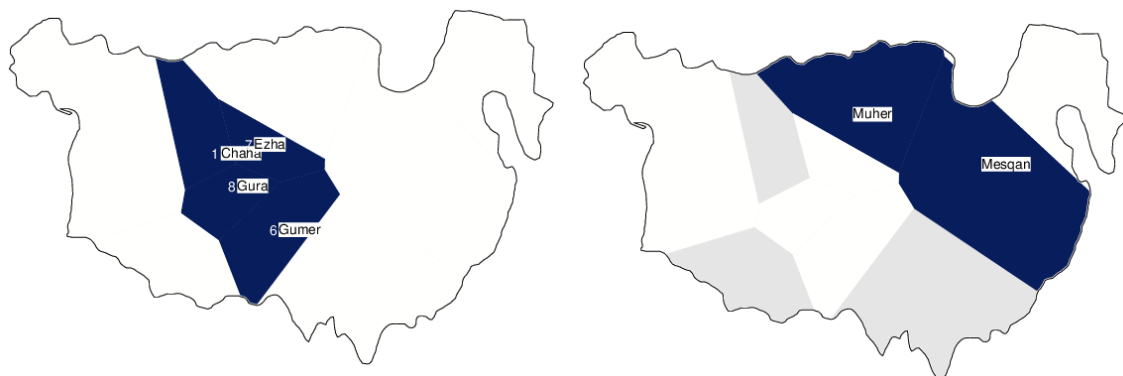
Table 7.2 presents the distinctive phonetic features of the ten top selected words of the West Gurage varieties. In the table, the pronunciation of each word in the group and outside the group are compared. The phonetic features which characterize the West Gurage languages are indicated using bold font. They are distinctive features since they are different in terms of their pronunciation in the West Gurage area as compared to the way they are pronounced elsewhere outside the West Gurage area. As can be seen from the table, the distinctive phonetic features that are specific to the West Gurage area include alveolar nasal [n], absence of velar fricative [x], velar plosive [g], geminated nasal bilabial [mm], close-mid front vowel [e], lateral approximant [l], alveolar trill [r] and geminated postalveolar fricative [ʃ]. These are just some of the phonetic features of the West Gurage area.



Table 7.2: Ten top ranked words of the West Gurage varieties

Rank	Words	in Cluster	Not in the Cluster	Characteristic Features
1	hen	[kuttə <b>n</b> a]	[kut(t)ə(a) <b>ra</b> ]	[n]
2	tiger	[z <b>o</b> gara]	[zə <b>g</b> <sup>w</sup> ərə(a)]/[zə <b>g</b> <sup>w</sup> arə]	[g]
3	hair	[gun <b>n</b> ən]	[gun(n)ə <b>r</b> ]/[gun <b>?</b> ər]	[n]
4	eight	[s <b>i</b> mmut]	[sunt]/[š(s) <b>i</b> mut]	[mm]
5	‘ingera’	[in <b>g</b> era]	[in <b>g</b> ərə]/[in <b>g</b> ər <b>?</b> ə]	[e]
6	string road	[wə <b>l</b> gəǵa]	[wə <b>n</b> gəǵa]/[wə <b>g</b> ənǵa]	[l]
7	wheat	[s <b>i</b> rre]	[s <b>i</b> rə <b>y</b> ]/[s <b>i</b> nə(e) <b>?</b> ]	[r]
8	red	[bi <b>š</b> ša]	[bi <b>š</b> (ə)]/[bu <b>š</b> ə(o)]	[šš]
9	egg	[ank' <sup>w</sup> ə]	[ink' <b>u</b> lə]/[ink' <sup>w</sup> <b>r</b> ə]	[∅]
10	this	[zi]	[ <b>ʒ</b> i <b>x</b> ]	[∅]

The geographical distribution of the characteristic features can also be illustrated using maps. Figure 7.2 displays the maps of the pronunciation of the word ‘cloud’ and ‘hen’ in the Central West Gurage and in the West Gurage areas respectively. In Figure 7.2 (a), the dark blue color shows the area where the pronunciation of the word ‘cloud’ is very distinct, compared to other areas. In the same manner, in Figure 7.2 (b), the dark blue color shows the area where the pronunciation of the word ‘hen’ is very peculiar.



(a) ‘cloud’- selected as a lexical item containing the most characteristic phonetic feature of the Central West Gurage dialects

(b) ‘hen’- selected as a lexical item containing one of the most characteristic phonetic features of the West Gurage dialects

Figure 7.1: Language areas based on the pronunciation of words (a) ‘cloud’ and (b) ‘hen’

## 7.1.2 Horizontal Diffusion of Linguistic Features

In 4.4.1.2.1, it was indicated that hierarchical clustering has its own disadvantages, i.e., it does not make a distinction whether the process of transmission of features is from a mother language variety to daughter varieties or due to horizontal diffusion. This section examines whether the classifications presented in Chapter six are the results of the genealogical relationship (vertical transmissions) among the language varieties or the reflections of the contact among the languages (horizontal transmissions). As discussed in 4.4.1.2.1, in the present study, two processes i.e., vertical and horizontal transmissions are differentiated using the Neighbor-net representation and lexicostatistical skewing.

### 7.1.2.1 Network Representation

It was indicated in 4.4.1.2.1 that one of the approaches towards distinguishing the horizontal transmission of linguistic features from the vertical transmission is using Neighbor-net representation. An important aspect of the Neighbor-net algorithm is that if the input distances are circular (due to horizontal transmission), it will return collections of circular splits, networks. In other words, it provides net-like classifications of languages. Nonetheless, if the distance inputs are additive, it will return a binary tree. This tendency of the Neighbor-net representation renders an opportunity to examine if the distance data is tree-like or net-like (Prokić et al., 2012). In other words, the Neighbor-net algorithm can be used to inspect whether the classification based on the lexical distance presented in Chapter six represents a genealogical relationship among the language varieties. The result of this analysis is illustrated in Figure 7.2. Figure 7.2 shows three broad groups of languages; one group consisting of {Gura, Chaha, Ezha and Gumer}, the second group consisting of {Muher, Mesqan, Silt'e and Kistane} and the third of {Edegagn and Inor}. The net-like structure of the Neighbor-net presentation shows the influence of language contact on the distance among the languages. The contact among the languages could be the result of geographical proximity. The classical example is the affinity between Kistane and Silt'e. As the two languages are genealogically only remotely

related, their close similarity must be largely due to the areal diffusion.

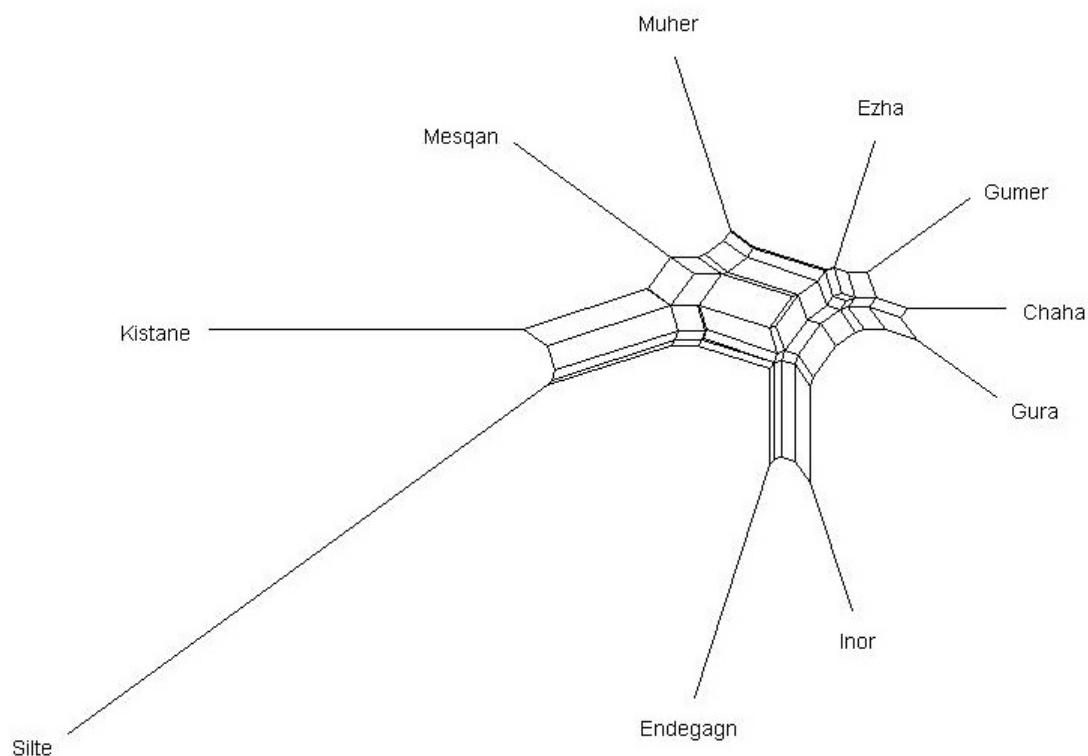


Figure 7.2: Neighbor-net dialects, Network-like structure of the data, based on lexical differences

### 7.1.2.2 Estimated Borrowed Features

This section, using lexicostatistical skewing, further examines to what extent the classification of the South Ethiosemitic languages presented in Chapter six are influenced by the borrowed phonetic and lexical features. Lexicostatistical skewing method (see §4.4.1.2.2) was applied to the combined classification which was presented in 6.4. In 4.4.1.2.2, the lexical skewing between  $L_A$  and  $L_B$  with respect to  $L_D$  was defined as follow.

$$S_{AD} - S_{BD}$$

Positive skewing shows borrowing between  $L_A$  and  $L_D$ . The negative skewing means many things as discussed in 4.4.1.2.2. Moreover, the aggregate skewing of  $L_A$  with respect to  $L_D$  was defined as the average of skewing between  $L_A$  and each of its sibilings with respect to  $L_D$ .

### 7.1.2.2.1 Estimated Borrowed Phonetic Features

In 6.4, the South Ethiosemitic languages were classified into five groups by combining the structural, functional and perceptual distance measures; Central West Gurage - {Chaha, Gura, Gumer, Ezha}, West Gurage - {Muher and Mesqan}, Peripheral West Gurage - {Endegagn, Inor}, {Silt'e} and {Kistane}. This section presents the percentage of borrowed phonetic features within the first three groups. Since Silt'e and Kistane are different from almost all other Gurage varieties, they are not further examined in this section. As the aim here is illustrating whether the similarity among the languages is due to borrowing, varieties that do not form group with other varieties are irrelevant. Table 7.3 and 7.4 present the percentage of phonetic features that are borrowed among the Central West, Peripheral West and West Gurage language varieties. The borrowed phonetic features were estimated based on the phonetic similarity index (see Appendix C.1). The phonetic similarity index was obtained by subtracting the phonetic distance from 100 ( $s_{\text{index}} = 100 - \text{phonetic distance}$ ). Figure 7.3 presents the aggregate phonetic skewing of the Central West Gurage language varieties with respect to the the Peripheral West and West Gurage varieties.

Table 7.3: Aggregate phonetic skewing of the Central West Gurage languages (CWG) with respect to Peripheral West Gurage (PWG) and West Gurage languages(WG). The numbers are percentages.

$\delta S_{CentralWest}^{PWG\&WG}$	PWG		WG	
	Endegagn	Inor	Mesqan	Muher
Chaha	0.3	1.7	1.3	0.7
Ezha	-1.0	-2.3	0.0	1.7
Gumer	0.3	-1.0	-1.3	-1.0
Gura	0.3	1.7	0.0	-1.0

Table 7.3 shows a small percentage of borrowed phonetic features (1.7%) between Inor and Chaha, Gura and Inor and Ezha and Muher. The percentage of borrowed phonetic features among the remaining language varieties is extremely small. As can be seen from the table, often the skewing results are negative. For instance, the negative skewing between Ezha and Inor (-2.3%) shows that probably Ezha might

have been influenced by other neighboring languages that may not have similar influence on other sister Central West Gurage language varieties. Table 7.3, in general, shows that the Central West Gurage languages borrowed only a small percentage of phonetic features from the West Gurage and Peripheral West Gurage varieties. Based on 4% skewing threshold set in 4.4.1.2.2, the amount of borrowed phonetic features is significantly small.

Table 7.4: Aggregate phonetic skewing of the Peripheral West and West Gurage with respect to other varieties

	CWG			WG		
$\delta S_{PeripheralWest}^{CWG\&WG}$	CH <sup>a</sup>	EZ	GM	GU	MS	MU
Endegagn	-6.0	-4.0	-4.0	-6.0	-3.0	-2.0
Inor	6.0	4.0	4.0	6.0	3.0	2.0
	CWG			PWG		
$\delta S_{WestGurage}^{CWG\&PWG}$	CH	EZ	GM	GU	EN	IN
Mesqan	3.0	1.0	2.0	3.0	1.0	2.0
Muher	-3.0	-1.0	-2.0	-3.0	-1.0	-2.0

<sup>a</sup>EN = Endegagn, IN = Inor, EZ = Ezha, GM = Gumer, GU = Gura, MS = Mesqan, MU = Muher and CH = Chaha, the results in percentage.

The first half of Table 7.4 shows the aggregate phonetic skewing of Peripheral West Gurage with respect to Central West Gurage and West Gurage. The table shows a high percentage of borrowed phonetic features between Inor and Chaha (6%), Inor and Ezha (4%), Inor and Gumer (4%) and Inor and Gura (6%). In general, there is a high percentage of borrowed phonetic features between Inor and all the Central West Gurage language varieties. The relationship between Endegagn and the Central West Gurage language varieties is quite the opposite. There is a negative skewing between Endegagn and Chaha (-6%), Endegagn and Ezha (-4%), Endegagn and Gumer (-4%) and Endegagn and Gura (-6%). The negative skewing between Endegagn and Central West language varieties could be due to the influence of the neighboring non-Semitic languages on Endegagn. The phonetic difference between Inor and Endegegn previously reported in 6.1 could also be due to either the contact between Inor and Central West Gurage languages or due to

the influence of other non-Semitic neighboring languages on Endegagn.

The second half of Table 7.4 presents aggregate phonetic skewing of the West Gurage languages with respect to the Central West and Peripheral West Gurage languages. The table shows some degree of positive skewing between Mesqan and Chaha (3%) and Mesqan and Gura (3%). In contrast to Mesqan, there is a negative skewing between Muher and Chaha (-3%), Muher and Gura (-3%). Nonetheless, the percentage of borrowed phonetic features is small in both positive and negative skewing cases.

### 7.1.2.2.2 Estimated Borrowed Lexical Items

The percentage of the borrowed lexical items among the ten South Ethiosemitic languages classified in 6.4 was also estimated using the lexical skewing. Table 7.5 and 7.6 present lexical skewing results which were computed based on the lexical similarity index (see Appendix C.2). The lexical similarity was the percentage of shared cognates among pairs of the language varieties. Table 7.5 presents the aggregate lexical skewing of the Central West Gurage languages with respect to the Peripheral West and West Gurage languages. The table shows a relatively high degree of negative lexical skewing between Gura and Muher (-4.0%) and Gumer and Muher (-3.7%). The lexical borrowing among other languages varieties extremely small. As can be seen from the table, there are also other negative skewing results but which are not significantly high.

Table 7.5: Aggregate lexical skewing of the Central West Gurage languages with respect to Peripheral West Gurage and West Gurage

$\delta S_{CentralWest}^{PWG\&WG}$	PWG		WG	
	Endegagn	Inor	Mesqan	Muher
Chaha	0.3	-1.0	-1.7	0.3
Ezha	1.7	-2.0	1.7	1.7
Gumer	-1.0	0.3	-1.7	-3.7
Gura	-1.0	2.0	-2.0	-4.0

The first part of Table 7.6 presents aggregate lexical skewing of the Peripheral West Gurage languages with respect to the Central West and West Gurage lan-

guages. The table shows a very high percentage of borrowed lexicons between Inor and Chaha (4%), Inor and Gumer (7%), Inor and Gura (7%) and Inor and Muher (5%). Contrary to these, there is a high percentage of negative skewing between Endegagn and Chaha (-4%), Endegagn and Gumer (-7%), Endegagn and Gura (-7%) and Endegagn and Muher (-5%). These results, in general, show that Inor has a strong contact with the Central West Gurage language varieties, contrary to Endegagn. It has also some degree of contact with West Gurage language varieties, especially with Muher. The negative lexical skewing between Endegagn and the Central West and West Gurage languages shows that Endegagn might have been affected by other neighboring languages, compared to Inor.

Table 7.6: Aggregate lexical skewing of the Peripheral West Gurage and West Gurage language varieties with respect to other varieties

	CWG			WG		
$\delta S_{PeripheralWest}^{CWG\&WG}$	CH <sup>a</sup>	EZ	GM	GU	MS	MU
Endegagn	-4.0	-2.0	-7.0	-7	-3.0	-5.0
Inor	4.0	2.0	7.0	7.0	3.0	5.0
	CWG			PWG		
$\delta S_{WestGurage}^{CWG\&PWG}$	CH	EZ	GM	GU	EN	IN
Mesqan	-3.0	-2.0	-1.0	0	3.0	0
Muher	3.0	2.0	1.0	0	-3.0	0

<sup>a</sup>EN = Endegagn, IN = Inor, EZ = Ezha, GM = Gumer, GU = Gura, MS = Mesqan, MU = Muher and CH = Chaha, the results in percentage.

The second half of Table 7.6 presents the aggregate lexical skewing of West Gurage languages with respect to the Central West and the Peripheral West Gurage languages. As can be seen from the table, both Mesqan and Muher have a marginal contact with both Central West and Peripheral West Gurage language varieties. The lexicostatistical skewing results, both based on phonetic and lexical similarity measures, indicate a strong tendency of borrowing between Inor and West Gurage language varieties.

### 7.1.2.3 The Influence of Local Influential Languages

As indicated in 2.1 and 2.2, the Gurage languages have been under the influence of Semitic, Cushitic and Omotic languages. It was also illustrated in 7.1.2.1 and 7.1.2.2 that the similarity or difference among the South Ethiopsemitic languages is partly the outcome of language contact. In 4.4.1.3, it was further indicated that the two most influential non-Gurage languages of the area are Amharic and Oromo. This section illustrates the degree of phonetic and lexical similarity between the Gurage varieties and these two influential languages. In line with the discussion in 5.4.1, the phonetic and lexical similarities were computed by comparing all the 240 lexical items using the Levenshtein algorithm.

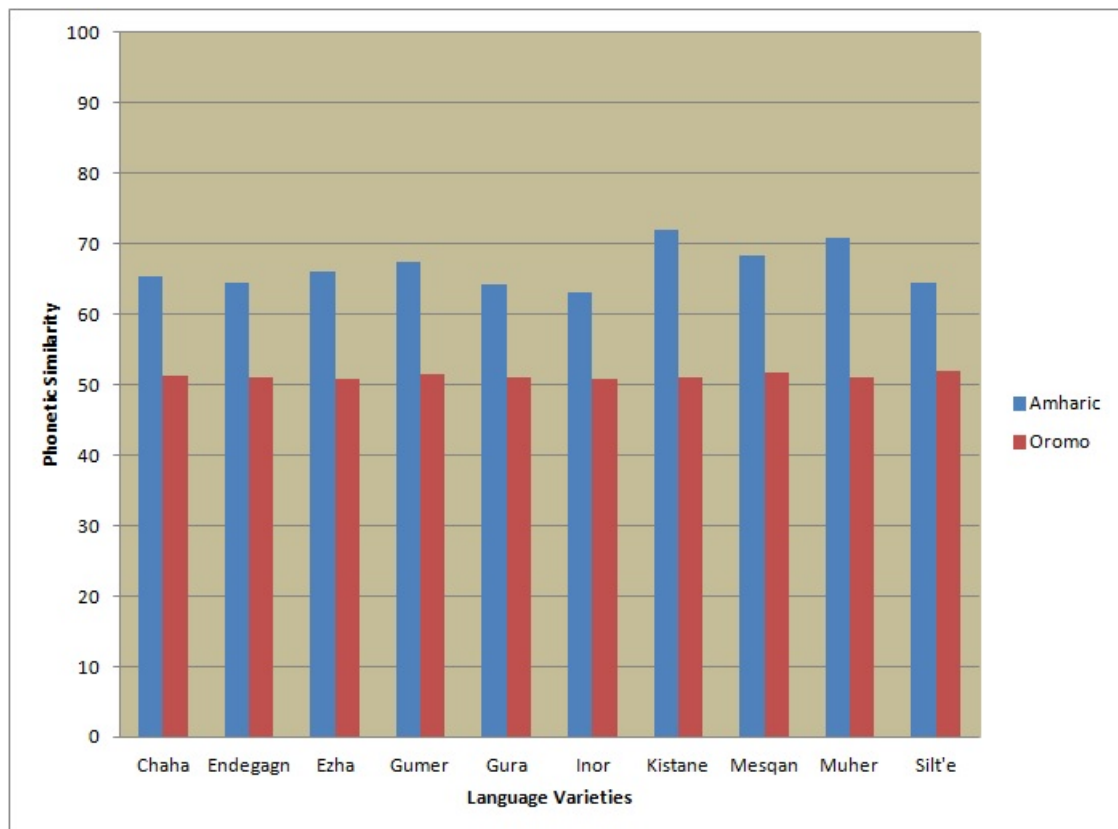


Figure 7.3: Percentage of shared phonetic features between the South Ethiopsemitic varieties and the two dominant languages

Figure 7.3 presents the percentage of phonetic features that are shared between Amharic and Oromo and the Gurage varieties. The Figure shows more than 60% phonetic similarity between Amharic and the Gurage varieties. Kistane, Mesqan



and Muher have relatively very large phonetic similarity with Amharic. Oromo shares about 50% phonetic similarity with almost all the Guraga varieties. There is little difference among the Gurage varieties in terms of the degree of their phonetic similarity with Oromo.

Figure 7.4 shows the percentage of shared lexicons between the two languages (Amharic and Oromo) and the Gurage varieties. The figure shows that many of the Gurage varieties share more than 50% of the lexical items with Amharic. Relatively, there is a high lexical similarity between Amharic and Kistane, Muher and Mesqan. The high lexical similarity between Amharic and the Gurage varieties could be due to both the genealogical relationship and the contact between Amharic and the Gurage varieties. Furthermore, each Gurage variety shares at least 20% of the Oromo lexical items. Compared to other Gurage varieties, Silt'e shares slightly higher number of lexicons with Oromo. This lexical similarity must be due to the long history of contact between the Gurage varieties and Oromo. Both the phonetic and lexical evidences indicate that Muher, Mesqan and Kistane have very close similarity with Amharic, compared to other Gurage varieties.

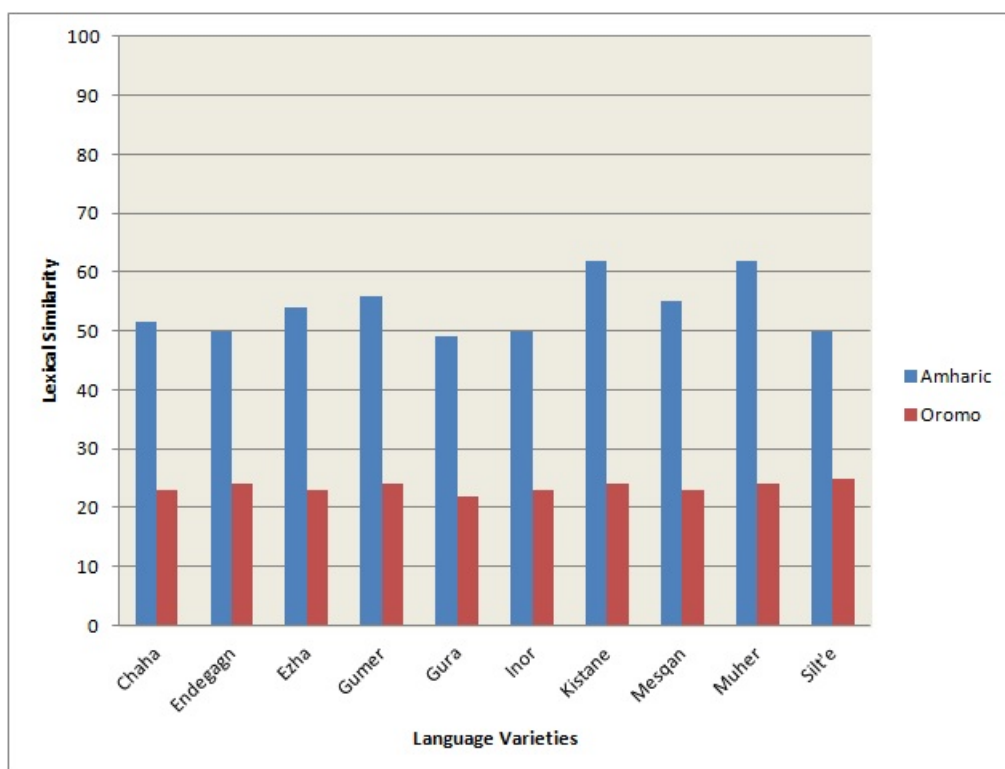


Figure 7.4: Percentage of shared lexicons between the South Ethiiosemitic varieties and the two dominant languages

## 7.2 Non-Linguistic Determinants

In this section, the results of the non-linguistic determinants that have contributed to the distance among the South Ethiosemitic languages are presented. As indicated in the introduction of this chapter, three major non-linguistic factors were examined: the contact among the speakers, geographical distance and population size, and language attitude.

### 7.2.1 Contact among the Speakers

The degree of contact among the speakers of the South Ethiosemitic languages was investigated using the background questionnaire (see Appendix A.1). The experimental subjects presented in 5.1.2 were asked four questions that are related to contact among the speakers: (1) what other languages they do speak other than their native language (Part II, Q.2); (2) which languages do their parents speak (Part II, Q.3); (3) what languages are frequently spoken by their close friends (Part II, Q.4); and (4) what languages are spoken in most of the schools they attended (Part II, Q.7). The participants reported contact with both Semitic (Amharic and Gurage varieties) non-Semitic languages: Oromo, Hadiyya and Welayta. The first two non-Semitic are Cushitic languages while the last one is Omotic language.

Figure 7.5 shows the number of languages<sup>18</sup> that the speakers of each language variety have contact with. The figure shows that many speakers of the Gurage varieties have exposure to Oromo. Moreover, Silt'e, Kistane and Chaha speakers have a notable contact with the speakers of several other Gurage varieties. Gura, Gumer and Inor have less exposure to the speakers of other Gurage varieties. It is important to mention here that some of the participants were not able to name all the Gurage varieties precisely. For example, with regard to the question 'what other languages do they speak other than their native languages?', a couple of participants

---

<sup>18</sup>Chaha speakers have contact with the speakers of eight languages, Gumer with the speakers of three languages, Ezha with six, Mesqan with six, Muher with five, Kistane with Eight, Silt'e with ten, Inor with four, Endegagn with six, Amharic with ten, Oromo with eight, Welayta with two and Hadiyya with one

answered ‘Sebat Bet Gurage’, which is a common name mainly for Chaha, Ezha, Gura, Gumer and Muher taken together. In some cases, there were also participants who simply responded ‘Gurage’ or ‘Guragigna’.

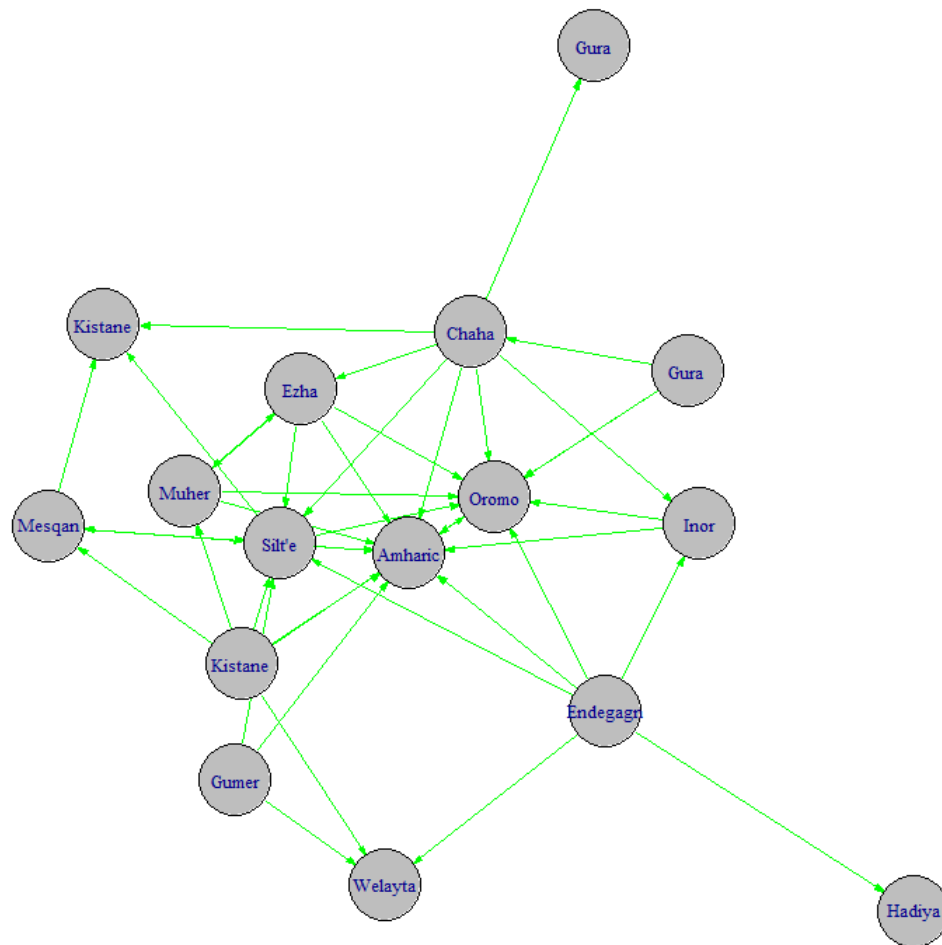


Figure 7.5: Contact of the speakers of the ten South Ethiosemitic languages.

Figure 7.5 further shows that most speakers of the Gurage varieties have exposure to Amharic. Indeed almost all Gurage speakers are bilinguals; they speak at least Amharic as a second language. This is not surprising since Amharic is the language of instruction in schools in the Gurage and Silt'e Zones. The figure also

shows the contact between Oromo speakers and Gurage speakers. Muher, Kistane, and Silt'e speakers are among those that have contact with the Oromo speakers. Readers should note that the arrow which connects Amharic and Oromo also connects Kistane and Oromo. Speakers of Endegagn have contact with the speakers of the Omotic language, Welayta. Endegagn speakers have also contact with the speakers of Hadiyya, another Cushitic languages. Most importantly, however, Figure 7.5 shows that the speakers of all Gurage varieties have exposure at least to some of other Gurage varieties. It also seems strange that one participant reported the contact between Kistane and Welayta (see Figure 7.6). As the two languages are geographically distant from each other, this must be because of the history of movement of a family member of the respondent. The contact between the speakers of Gumer and Welayta must also be due to movement since the two languages are not geographically adjacent to each other.

It can also be seen from the figure that several Gurage varieties have contact with Chaha and Silt'e speakers. This could be due to several factors, but mainly due to the geographical locations of the language areas and the history of the languages. For instance, there is a main road which connects Kistane, Mesqan and Silt'e areas. This could facilitate the interaction or communication among the speakers of the three Gurage varieties. Furthermore, historically, Chaha has been the language of traditional courting (Yejoka) and language of teaching Bible. This could enhance the contact between Chaha speakers and the speakers of other Gurage varieties. The contact between the speakers of Kistane and Oromo and that of Silt'e and Oromo could also be mainly due to the geographical proximity of Kistane and Silt'e; the two areas are found at the border of the Oromiya Region where Oromo is spoken (see Figure 2.2). The issue of geographical distance will be discussed in the next section.

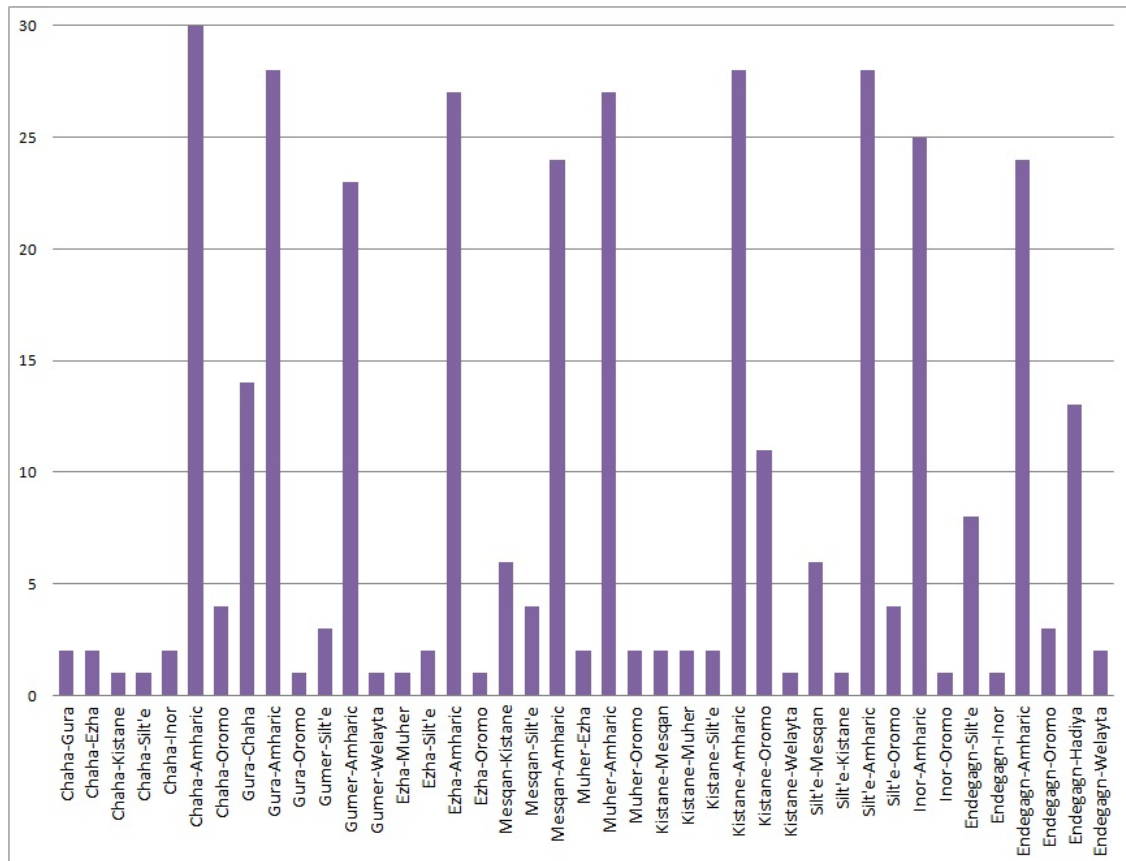


Figure 7.6: Degree of contact among the speakers of the ten South Ethiosemitic languages

Figure 7.6 shows the number of speakers (of a language variety) that have contact with the speakers of other languages. The figure shows that many of the speakers of each Gurage variety have contact with Amharic (at least 24 out of 30 subjects in each group). Moreover, many Gura speakers have contact with the speakers of Chaha. This is not surprising since both Chaha and Gura are spoken in Chaha area (see Figure 2.2). More importantly, several Kistane speakers have contact with the speakers of Oromo. As indicated above, this could be due to the geographical proximity between the speakers of the two languages, and to the main road that extends from Addis Ababa to Oromiya Region and Kistane then to the remaining southern part of Ethiopia. Many speakers of Silt'e have also contact with the speakers of Mesqan. This is also not surprising since the two languages are spoken adjacent to each other, and they are connected by the main road that crosses the two language areas on the way from Addis Ababa to the rest of the southern part of Ethiopia. Silt'e speakers have also some degree of contact with the speakers of Oromo. Only

some speakers of Endegagn have a contact with the speakers of Hadiyya.

### 7.2.2 Geographical Distance and Population Size

It was indicated above that population size and geographical distance are some of the factors that determine the linguistic distance. As discussed in 4.4.2.1, the Gravity model proposes that the linguistic similarity between any given pairs of languages is inversely proportional to the geographical distance, but directly proportional to the the population size. To examine this claim in the context of the South Ethiosemitic languages, two types of geographical distance measures were employed: the driving time between each site (in minute) and the driving distance between the sites (in km). The phonetic and lexical similarities among the language varieties were used as the similarity indices (cf: 5.4.2). The similarity index was computed by subtracting the linguistic distance from 100. It is also important to mention that Gura was not included in this analysis since it is spoken in the Chaha district, and there is no data available about the population size of the speakers of Gura. For the remaining nine varieties, the population size (see §2.2.3) was taken from Ethiopian National Census (2007). The detailed description of the source of the population size was presented in section 5.4.2.

The prediction of the Gravity model based on the driving distance (km) and the phonetic and lexical similarity indices is presented in Table 7.7. Driving distance is the distance of the main roads that connect the language areas. The correlation coefficient of Pearson's correlation was computed both for the distance between the language areas, and the square of the distance between the language ares (based on the prediction of Trudgill's Gravity model).

Table 7.7: The influence of population size and geographical distance

	Phonetic Similarity	Lexical Similarity
Product <sup>a</sup> of Population Size	-.461	-.721
Distance in km	-.600	-.637
Square of Distance in km	-.631	-.671

<sup>a</sup>The product of population size means the size of the population of one area multiplied by the size of the population of another area. The distance between Gura and other Gurage varieties was not included in this analysis since Gura is spoken in the Chaha district and that the exact number of the Gura speakers is not known.

As can be seen from the table, there is an inverse relationship (negative correlation coefficient) between the product of population size and the phonetic similarity,  $r = -.461$ . Contrary to the prediction of the Gravity model, the table shows that as the size of population increases, the phonetic similarity between pairs of languages decreases. The same is true for the lexical similarity. As the population size increases, the lexical similarity decreases,  $r = -.721$ . On the other hand, Table 7.7 shows that there is also an inverse relationship between the geographical distance and the linguistic similarity. This is true both for the phonetic and lexical similarities. This means that as the geographical distance between two sites increases, the phonetic and lexical similarity between pairs of languages decreases. This is the case both in the correlation between linguistic similarity (lexical and phonetic) and geographical distance, and the linguistic similarity and the square of the geographical distance. In general, the correlation between the linguistic similarity and the geographical distance supports the Gravity model, whereas the correlation between the linguistic similarity and the population size is counter evidence against the Gravity model.

Table 7.8 presents the result of the correlation coefficients of the two linguistic distances (phonetic and lexical) and the geographical distance which was measured based on the driving time. The driving time is the average time that public transportation, mainly buses, takes to travel from one language site to another. The correlation coefficient was computed both for the distance between the sites and the

square of the distance between the sites. As can be seen from the table, there is a negative correlation between the phonetic similarity and the population size,  $r = -.461$ . In the same manner, there is an inverse relationship between lexical similarity and population size,  $r = -.721$ .

Table 7.8: Correlation coefficients of phonetic similarity, geographical distance and population size

	Phonetic Similarity	Lexical Similarity
Product <sup>a</sup> of Population Size	-.461	-.721
Distance in min	-.651	-.610
Square of Distance in min	-.650	-.613

<sup>a</sup>Product of population size means the size of population of one area multiplied by the size of population another area. The distance between Gura and other Gurage varieties was not included in this result since Gura is spoken in the Chaha district and the exact number of Gura speakers is not known.

In other words, similar to the results presented in Table 7.7, the results presented in the Table 7.8 show that there is no positive correlation between the population size and the degree of linguistic similarity. Rather, contrary to the prediction of the Gravity model, there is an inverse relationship between the population size and the linguistic similarity. It is strange that the correlation between the two is negative which means that the smaller the population size the higher the linguistic similarity. On the other hand, precisely as predicted by the Gravity model, there is very strong inverse correlation between the geographical distance and the linguistic similarity, both phonetic and lexical. In general, the results presented in Table 7.7 and 7.8 do not confirm the social gravity impact claimed by the Gravity model.

It could be the cases that while geographical proximity is a necessary condition for the diffusion of linguistic features, the impact of the population size is later determined by other several non-linguistic factors. The absence of correlation between the population size and linguistic similarity can also be due to the population heterogeneity in each of the language area. In each Gurage district, there are many non-Gurage communities. Moreover, the absence of correlation between the population size and the linguistic similarity could be the result of the sociolinguistic reality



in the Gurage area. As presented in 2.2.2, the Gurage communities are bilingual. Amharic is a dominant language in towns in the Gurage area. Therefore, may be, it is the case that the diffusion of Amharic phonetic and lexical features from one town to another has overtaken the diffusion of the features of the Gurage varieties. Hence, in order to exclude the influence of Amharic in the urban areas (towns), Pearson's correlation was computed only between the rural population and the linguistic similarity. Nonetheless, the correlation between the two remains negative; the correlation between the rural population and phonetic similarity,  $r = -.466$ , and the correlation between the rural population and lexical similarity,  $r = -.724$ .

The correlation coefficients between the two geographical distance measures (driving distance and driving time) and the two structural distance measures (phonetic and lexical similarities) were also compared to determine whether it is the driving distance or the driving time that is a better measure of geographical distance. The results of this analysis shows that the correlation between phonetic similarity and driving distance and that of the phonetic similarity and driving time are not significantly different; Hotelling's t-test,  $t = .810$ ,  $p = .424$ . Likewise, the correlation between lexical similarity and driving distance, and that of the lexical similarity and driving time are not statistically significant; Hotelling's t-test,  $t = -.424$ ,  $p = .674$ . These results, in general, show that the time required to travel from one language area to another and the distance between each of the language area have similar impact on the frequency of contact among the speakers of the languages.

### 7.2.3 The Attitude of the Speakers

In 4.4.2.2, it was indicated that the speakers' attitude is one of the factors that may determine the distance among related languages. This section presents the results of the relationship between the speakers' attitude and the distance among the language varieties. The attitude of the speakers was measured using a Likert scale that ranged from *not beautiful* (0) to *beautiful* (10), based on the speakers' response after listening to the recordings of a fable 'The North Wind and the Sun' (see §5.3.2 for details). Table 7.9 shows a strong negative correlation between the

speakers' attitude and the two structural distances; between the speakers' attitude and phonetic distance ( $r = -.751$ ) and between the speakers' attitude and lexical distance ( $r = -.705$ ). This means that the lesser the linguistic distance between the language varieties, the more positive the attitude of the speakers towards the languages. In other words, the speakers are more positive about the languages that they are familiar with. In the same manner, the table shows that the speakers have positive attitude towards the languages they understand better (there is a strong negative correlation between the speakers' attitude and the functional distance,  $r = -.682$ ). Finally, there is a very strong negative correlation between the speakers' attitude and perceptual distance,  $r = -.959$ : the participants are more positive to the languages they believe similar to their native language.

Table 7.9: Correlations between speakers' attitude and linguistic distance

		Attitude
Structural	Phonetic	-.751
	Lexical	-.705
Functional		-.682
Perceptual		-.959

Table 7.10 compares the correlation coefficients presented in Table 7.9. The table shows that there is no a significant difference between the correlation coefficient of the lexical distance and the speakers' attitude, and the phonetic distance and speakers' attitude. In the same manner, there is no statistically significant difference between the correlation coefficient of phonetic distance and speakers' attitude and that of the functional distance and speakers' attitude. Likewise, there is no a significant difference between the correlation coefficient of the lexical distance and speakers' attitude and that of the functional distance and speakers' attitude.

Nonetheless, Table 7.10 indicates that the perceptual distance is more strongly associated with language attitude, as compared to the phonetic distance;  $-.959 > -.751$ ,  $r = -4.390$ ,  $p < .001$ , Fisher's z-transformation. In the same manner, the correlation coefficient of the perceptual distance and the speakers' attitude ( $-.959$ ) is significantly greater than the correlation coefficient of the lexical distance and the

speakers' attitude (-.705),  $z = -4.840$ ,  $p < .001$ , Fisher's z-transformation. Moreover, the correlation coefficient of perceptual distance and the speakers' attitude (-.959) is significantly greater than the correlation coefficient of functional distance and speakers' attitude (-.682);  $t = 8.758$   $p < .001$ , Hotelling's t-test.

Table 7.10: Correlations among the three dimensions of distance and language attitude

Compared Coefficients <sup>a</sup>	t/z values	p.value	Test
$r_{PD\text{AT}}$ vs. $r_{LD\text{AT}}$	-0.450	.653	Fisher's z-transformation
$r_{PD\text{AT}}$ vs. $r_{FD\text{AT}}$	-0.653	.514	Fisher's z-transformation
$r_{PD\text{AT}}$ vs. $r_{PcpD\text{AT}}$	-4.390	.001	Fisher's z-transformation
$r_{LD\text{AT}}$ vs. $r_{FD\text{AT}}$	-0.231	.839	Fisher's z-transformation
$r_{LD\text{AT}}$ vs. $r_{PcpD\text{AT}}$	-4.840	.001	Fisher's z-transformation
$r_{FD\text{AT}}$ vs. $r_{PcpD\text{AT}}$	8.758	.001	Hotelling's t-test

<sup>a</sup> $r$  = correlation coefficient, PD = Phonetic Distance, AT = Language Attitude, LD = Lexical Distance, FD = Functional Distance, PcpD = Perceptual Distance

The results presented in Table 7.10 depict three important points. First, all the distance measures strongly correlate with the attitude of the speakers. This shows a general tendency of positive relationship between language similarity and the attitude of the speakers. The general tendency is that as the linguistic similarity between languages increases, the speakers' positive perception toward the languages also increases. The second important point is that there is a weak connection between the speakers' attitude and the functional distance as compared to the perceptual distance. In other words, it is not always the case that the speakers' actual understanding of a language is dependent on the attitude of the speakers towards that language. More importantly, perceptual distance is highly sensitive to the attitude of the speakers. Among all the distance measures, it is the perceptual distance that is most likely to be affected by the speakers' attitude.

### 7.3 Discussions

The aim of this chapter was identifying the linguistic and non-linguistic determinants that determine the linguistic distance among the South Ethiosemitic languages. Three types of linguistic determinants were examined, i.e., characteristic features (shibboleths), diffusion of features due to contact and the influences of Amharic and Oromo. With regard to the characteristic features, special phonetic features that characterize two language areas, the Central West Gurage and West Gurage were identified. Based on the analysis of the most frequent top ranked words, the main characteristic phonetic features that make the Central West Gurage language varieties distinct from other Gurage varieties include labialized velar fricative [x<sup>w</sup>], velar ejective [k'], alveolar stop [t], alveolar nasal [n], palatal ejective [c'], labialized nasal bilabial [m<sup>w</sup>], bilabial plosive [b] and alveolar trill [r]. The distinctive phonetic features that are specific to the West Gurage area include alveolar nasal [n], lack of velar fricative [x], velar plosive [g], geminated nasal bilabial [mm], close-mid front vowel [e], lateral approximant [l], alveolar trill [r] and geminated postalveolar fricative [š]. These results show that the dialectometric approach to language variation can be applied not only to the aggregate linguistic data, but also to distinguish specific linguistic features which underpin the groupings of related languages - as was previously argued by Prokić et al. (2012), Nerbonne et al. (2011) and Wieling & Nerbonne (2011).

Determining the influence of contact-induced linguistic features on the distance among the South Ethiosemitic languages was another focus of this chapter. In this regard, the diagnosis using Neighbor-net algorithm shows that the lexical classifications of the South Ethiosemitic languages which was presented in Chapter six are influenced by the contact among the languages. The net-like structure of the Neighbor-net representation, and the affinity between, for example, Silt'e and Kistane which are otherwise historically only remotely related are taken as evidences to argue that language contact has contributed to the similarity among South Ethiosemitic languages. In the same manner, the estimation of borrowed phonetic and lexical features using lexicostatistical skewing shows that there is a high percentage of bor-

rowed phonetic and lexical features especially between Inor and Central West Gurage languages (Chaha, Gura, Gumer and Ezha), and some degree of contact between Inor and West Gurage languages (Muher and Mesqan). It seems that the borrowings are partly the consequences of the geographical adjacency among the languages. As can be seen from Figure 2.2, Inor area is bordered by Gumer in southeast and by Chaha (which has strong contact with Gura and Ezha) in west.

The borrowings among the languages may also be connected to other social and historical facts. For example, the high percentage of borrowed phonetic and lexical features between Inor and Central West Gurage languages could be due to the historical movement of the Gurage people from east to west as indicated by Menuta (2015). There is also the possibility of the influences of non-Semitic neighboring languages which might have caused language divergence. The negative phonetic and lexical skewing between Endegagn and Central West Gurage as well as the West Gurage languages could be due to the influence of non-Semitic neighboring languages on Endegagn since Endegagn speakers have contact with non-Semitic languages such as Hadiyya. In general, there are both phonetic and lexical signals which show that borrowing is one of the causes of the similarity among the language varieties.

The influence of Amharic and Oromo was also illustrated by examining the possible degree of borrowing between the two languages and the Gurage varieties. There is indeed borrowing between these two dominant languages and the Gurage varieties. Amharic shares more than 60% of the phonetic and more than 50% lexical features with many of the South Ethiosemitic language varieties investigated in the present study. Mesqan, Kestane and Muher are among the varieties that have been significantly influenced by Amharic. Oromo shares more than around 50% phonetic features and more than 20% lexical features with the Gurage varieties. The influence of Oromo seems stronger on Silt'e.

With regard to the non-linguistic determinants, results of the comparisons of geographical distance and linguistic distance substantiate the fact that the similarity among the ten South Ethiosemitic languages is mainly because of the geographical proximity among the languages. Adjacent languages are more similar than languages that are far apart. This is precisely the view of the Gravity model (Trudgill,

1974; Nerbonne et al., 2005). Nevertheless, contrary to the Gravity model, the population size does not have an impact on the linguistic distance among the languages. According to the Gravity model, population size is supposed to be a driving force behind the spread of linguistic innovation (Nerbonne et al., 2005). It assumes that bigger centers always affect smaller ones. The analyses performed both on the phonetic and the lexical similarity indices do not confirm this hypothesis. Similar results were reported in Nerbonne et al. (2005). There are also several other studies that could not replicate the prediction of the Gravity model (e.g., Boberg, 2000; Horvath & Horvath, 2001).

It can be the case that population size by itself is not the main determinant factor but other factors associated with it. For example, the impact of population size may become more prominent when the speakers of the languages are also economically and politically dominant. In the Gurage area, in general, there is no economically dominant social group. None of the varieties is politically dominant either since none of the varieties has been used for administration, schooling and media (except Silt'e in the Silt'e Zone). Chaha is relatively a privileged variety since it is occasionally used for religious teaching and in traditional court (Yejoka). However, it is less likely that it influences other languages with a significant magnitude. Boberg (2000) reported the negative influence of political borders on the diffusion of linguistic features between two big centers. However, this cannot be the case in the context of Gurage area since all the language varieties are spoken in the same country and even in a small geographical area. Another factor could be a negative attitude deriving from the association of a language with foreign rule as reported in Pavlenko (2006). Given that the Gurage people, in general, are positive about the Gurage varieties that they are familiar with, it is less likely that language imposition is the influential factor. Given the linguistic and social diversity in the Gurage area, this heterogeneity of the community can be another factor influencing the diffusion of linguistic features.

The influence of Amharic could be another plausible factor. Since all Gurage varieties are under the influence of Amharic and most speakers of Gurage varieties are speakers of Amharic, probably the potential influence of Gurage varieties with

a large population size has been suppressed by the dominance of Amharic. For a century, Amharic has remained the most influential language in the Gurage area. Speakers of every Gurage variety investigated in the present study have exposure to Amharic. Although Amharic was initially exported to the Gurage land through national enforcement, forced settlement, and the mobile nature of the Gurage community, it gradually became the most preferred language in the Gurage area. It is also recognized as the language of elites, business and politics. Nowadays, according to Menuta (2015), many Gurage people have positive attitudes towards Amharic. Indeed, many of them prefer Amharic to their local languages for schooling and administration.

In 7.2.3, it was also illustrated that there is a strong positive correlation between the similarity among the languages and the attitude of the speakers. The speakers are more positive about the languages they are familiar with. The negative attitude towards unfamiliar languages, to a certain extent, may show the conservative position of the Gurage community towards new languages. This may emanate from a conservative tribe-based tradition suggested by Hetzron (1972). The conservative culture may also be associated with the effort to preserve local identity. There is a strong correlation between the speakers' attitude and perceptual distance, as compared to the correlation between the structural distances and the speakers' attitude. This result has an important methodological implication. It shows that the use of the perceptual distance as a measure of the linguistic distance can sometimes be problematic since it can be influenced by the attitude of the speakers. Similar observations were previously made by Abu-Rabia (1996), Abu-Rabia (1998) and Golubović & Sokolić (2013).

In general, the factors examined in this chapter show that geographical distance and contact among the languages are the main determinants that have affected the linguistic distance among the South Ethiosemitic languages. Strong evidence of diffusion of linguistic features show that the classifications of the South Ethiosemitic languages (based on phonetic and lexical distance) which were presented in Chapter six reflect more of the areal similarity rather than the genealogical relationship among the languages. In this regard, the close similarity between the combined clas-

sification and the classification by Hetzron (1972) has an important methodological implication with regard to the classification of South Ethiosemitic languages. The classification of Hetzron (1972) has been considered as a genealogical classification of the South Ethiosemitic languages. Given that the classification of Hetzron (1972) is similar with the areal classifications reported in Chapter six, there is still a possibility that both the combined classifications and the classifications by Hetzron (1972) reflect the areal classifications, not a pure genealogical relationship among the languages as many scholars of Ethiosemitic languages think. Re-constructing the genealogical relationship among the languages is out of the scope of the present study. Hopefully, it would be an enterprise of future studies.



---

## Chapter 8

# Theoretical Implications and Future Directions

In this chapter, the methodological and theoretical implications of the results reported in Chapter six and seven are presented. The chapter first summarizes the linguistic and social complexity of the language area in which the present study was conducted (§8.1). Then, it projects the results reported in Chapter six and seven onto the complex linguistic and social reality of the Gurage areas to reflect on some of the theoretical issues that have remained contentious in dialectology. Among the debatable issues, the notion of the Gravity model will be raised and the controversy surrounding the role of population size and geographical distance in determining the linguistic distance among related languages will be discussed (§8.2). The long-standing controversy between Tree and Wave models will be discussed in light of the linguistic reality in the Gurage area (§8.3). The chapter also re-examines the approaches of the previous studies conducted on the classifications of the Gurage varieties in light of these controversies. It also summarizes issues related to the reliability of the perceptual distance, and proposes contexts in which the perception of the non-linguists can be best utilized to determine the linguistic distance among related languages (§8.4). The relationship between linguistic distance and language attitude will be summarized to illustrate contexts in which the attitude of the speakers should be taken into account in the process of measuring the distance among related languages (§8.5). The chapter concludes the present study by proposing

a combined approach towards measuring linguistic distance as one of the perspectives that future studies on linguistic distance should take into account (§8.6). The chapter does not aim to offer a full-fledged description of each of these broad and debatable issues. It rather aims to instigate more discussion on the matters and inspire further inspection and rigorous future investigations.

## 8.1 Intense Contact, Considerable Heterogeneity

The intense contact among the South Ethiosemitic languages on the one hand and the preservation of the linguistic diversity on the other hand has remained a paradox for the scholars of Ethiosemitic languages. There is a long history of contact among the language varieties (see §2.2.2), and even today there is an intense contact among the speakers of South Ethiosemitic languages (see §7.2.1). There are also shared phonetic and lexical features among some of the languages (see §7.1). Regardless of all this, there is an incredible linguistic diversity in the Gurage area. The co-existence of these two contradicting phenomena has remained a mystery. The factors that underpin this paradox could be both linguistic and extra-linguistic, which is largely overlooked in the previous studies. There are just a few works on the contact among the Gurage speakers (e.g., Menuta, 2015), and the contact between Gurage varieties and neighboring non-Semitic languages (e.g., Leslau, 1952; Meyer, 2005).

This is not to imply that the diversity in the Gurage area has not received any attention. Indeed there are scholars who were concerned about the sociolinguistic heterogeneity of this area. For instance, Hetzron (1972) suggests that clan-based social structure is one of the contributing factors. The Gurage communities are deeply divided based on language, religion and ethnicity (see §2.2.3). These diversities existed for centuries, and they are still very visible. There are two assumption about sources of ethnic or clan-based identity and social structure: *intrinsic* and *situational*. According to Eifert et al. (2010, p. 494), the intrinsic hypothesis assumes that ethnic identities are hardwired and they are intrinsically parts of who the people are, and their salience emanates from their link to the peoples' natural makeup. On the contrary, according to situational assumption, ethnicity is salient because of its

functionality; it serves as a useful tool for mobilizing people and building coalitions that can be deployed in the struggle for power and scarce resources. It appears to me that while the intrinsic assumption may best explain the source of ethnic and social diversity, it is the situational assumption that best explains the preservation or the sustainability of ethnic/tribal social organization. With regard to the Gurage communities, it was indicated in 2.2.2 that they are religiously divided because of periods dominated by Muslim and Christian invaders at different times in history. It was also indicated that there were continuous wars among local warlords. These conflicts in one way or another are connected to a fierce power and resource competition. What is more, the current political setup in the country accelerates such a competition; for instance, the current Ethiopian constitution divides districts in the Gurage area based on the ethnic background of the community. This clan-based administrative division might have fueled the power and resource competition that had already been preserved for several centuries. The administrative separation of Silt'e from the rest of the Gurage communities in 2001 is an illustration of such an influence. According to Meyer (2018), the motive of the separation was more economic than linguistic.

The competition for power and resources can also serve as glue to strengthen internal clan unity. For instance, in tribal community, each clan member assumes the responsibility of protecting the norms of the clan which includes safeguarding the clan's language identity. Deviation from an established social and linguistic norm becomes a betrayal of one's own essential identity, and is often punishable. There are various forms of traditional social and political Gurage organizations that keep within-clan cohesion alive. As extensively discussed by Shack (1966) and Zewde (2002), there are different forms of traditional legal systems in place which enforce this social responsibilities. As discussed in 2.2.3, there has been traditional system of administration in Gurage communities which manifests itself in three level of authority: village, clan and region. These legal systems promulgate different sets of laws such as *Sabugnat*, *Agar* and *Yajoka and Gordanna Sera*. *Sabugnat's* responsibilities include providing gifts for funerals and for weddings of first-born son, ensuring equal use of grass lands, and adjudicating in cases of dispute. *Agar* administers churches

and communal lands, following the rules laid down by the Gordanna Sera. Yajoka and Gordanna's responsibilities include defending the community against external aggression, enacting administrative and judiciary laws and supervising the implementations of the fundamental rules of law. It is based on these different levels of laws that the political, economic and social unity is sustained. Within-clan cohesion which is enforced by these legal processes may also have far-reaching linguistic consequences. Given that language is a strong identity marker among the Gurage communities, there is a certain way of using a language that the clan norm supports. Adherence to this specific linguistic norm, for example manner of speech, can gradually lead to language divergence among the various clans that constitute the different Gurage communities. Moreover, since children are also members of the community (clan), they are shaped by social rules and norms which include early language supervision and correction that could eventually maximize language divergence among the Gurage clans.

The fact that the Gurage people are mobile (see §2.2.2) may have effects on the linguistic and social diversity of the Gurage area. The economic and political interest of the Gurage people is not limited to the tiny Gurage area. Rich Gurage businessmen whose economic interests extend far beyond the Gurage land are usually not permanent residents of the Gurage areas, hence, they play insignificant roles in the social, economic and political harmonization of the area. Probably this has reduced the economic and political domination among the Gurages, which could have a profound sociolinguistic consequence, and rather created a situation in which the influence of one language on another is not strong enough to lead to assimilation among the Gurage varieties.

The current language policy of the country may also have its own contribution. As indicated above, the current Ethiopian constitution allows every ethnic group to develop and use its own language. Though this seems positive for the minority languages in general, it has instigated a strong sense of competition among the speakers of different Gurage varieties. Since each Gurage clan is committed to preserve its own language identity for the reasons outlined above, standardizing the Gurage varieties and using the for schooling and administration has remained

a challenge (Menuta, 2015; Meyer, 2018). Several standardization attempts have failed because of the lack of consensus among the Gurage communities and the political elites. As an option to ease inter-clan tension, Amharic has been used as the language of schooling and administration. It appears to me that community leaders and Gurage political elites are comfortable with using Amharic so as to ease the political tensions that may emerge otherwise.

## 8.2 The Simplicity of the Gravity Model

The basic notion of Gravity model (Trudgill, 1974) is that population size and geographical distance have an impact on the diffusion of linguistic features. Trudgill's Gravity model predicts the spread of linguistic innovations via social contact which is naturally promoted by proximity and population size. It further assumes that the relationship between the linguistic similarity and the geographical distance is quadratic, there may be a rapid diffusion at the center which gradually decreases as the geographical distance increases quadratically (see §7.2.2).

The present study confirmed the direct relationship between the linguistic distance and the geographical distance claimed by the Gravity model. Several previous studies also reported similar results (e.g., Alewijnse et al., 2007; Gooskens, 2005; Prokić, 2007; Nerbonne et al., 2005; Wieling et al., 2007). However, it is essential to state here that, the relationship between linguistic similarity and geographical distance can be sub-linear, not quadratic as projected by the Gravity model. The sub-linearity of the relationship between linguistic similarity and geographical distance has been noticed by several previous studies (e.g., Alewijnse et al., 2007; Prokić, 2007; Nerbonne et al., 2005; Wieling et al., 2007). In general, the nature of relationship between linguistic distance and geographical distance requires further investigations. It seems though that there is a general consensus among scholars about the fact that geographical distance is a major factor that determines the distance among related languages.

On the contrary, there is a debate about the relationship between the linguistic distance and population size (see §4.4.3.2 and 7.2.2). The present study could not

confirm the effect of the population size claimed by the Gravity model. According to the results reported in 7.2.2, there is even an extreme case where the relationship between population size and linguistic similarity are inversely related, which is quite the opposite of the hypotheses of the Gravity model. As discussed in 7.2.2, several previous studies could not confirm the notion of the social gravity either. Therefore, it may be the case that the notion of social gravity proposed by Trudgill (1974) is too general; it neglects several factors that determine the influence of the population size on the linguistic distance. For instance, the influence of social gravity is probably prominent only when the center is big enough (metropolitan cities) to have an influence on other centers. This was the case, for example, in Heeringa et al. (2011). Contrary to Nerbonne et al. (2005) which investigated the influence of social gravity in Low Saxon dialects, Heeringa et al. (2011) reported the influence of the population size after including Dutch language areas with high population size mainly big cities in the Netherlands. Much related to this is the issue of population density. The influence of metropolitan cities where the population density is very high can have a significant effect since the the population density may foster rapid diffusion of linguistic features. In other words, as also pointed out by Trudgill (1974, p.221), what really matters may not be just the size of the population but also the density of the population.

The Gravity model also neglects several social and linguistic factors presented in 8.1, and others which some of them are later discussed in Trudgill (1983) and Trudgill (2004). Absence of correlation between the linguistic similarity and the size of the centers in the Gurage area can be, for example, due to the complexity of social organization in the area. In principle, the social gravity assumes a genuine need for communication and frequency of communication as a base for the radiation of linguistic innovations (see §4.4.4.2 & 7.2.2). However, the frequency of conversation exchange among speakers and the degree of accommodation can be affected by numerous variables. As indicated in 8.1, tensions between different clans, for example, can minimize the frequency of contact. Moreover, though it is not the case in the Gurage area, the political and economic superiority of a minority group can lead to the influence of a minority language on the language of the majority. In other

words, contrary to the prediction of the Gravity model, the direction of influence can be non-hierarchical. Indeed, in many countries, the so called standard dialects are spoken by the socially, politically and economically dominant higher class small group of the society. In this way, the political, economic, and social status of the speakers favors the diffusion and expansion of the features of the standard dialects (Milroy & Milroy, 2012; Scassa, 1994).

The diffusion of linguistic features may also require dynamic economic and social settings. If the economic and social mobility is low, the interaction between the speakers of different language varieties can be minimal. A lingua franca can also minimize the influence of locally dominant languages. This is probably the case in the Gurage area; the influence of Amharic (the lingua franca of the country) might have suppressed the influence of local big centers on the small centers. In such a situation, the lingua franca may reap all the benefits which the acquisition of the local varieties could provide and exclude them from competitions. The conservative social tradition discussed in 8.1 (see §2.2.2) can be another determining factor. A community with conservative traditions and strong sense of self-identity may not be easily influenced by other languages in contrast to a community with vibrant and fluid social traditions. As the diffusion of linguistic features require free interactions among the speakers of different language varieties, the interaction could be hindered by such a conservative tradition. Especially if there is social marginalization and stereotyping associated with the clan identity, it could be an obstacle to cross-clan communication. A subjugated and marginalized community is very likely to be resistant towards the expansion of the dominant and oppressive languages of higher classes (see Abu-Rabia, 1996; Abu-Rabia, 1998; Golubović & Sokolić, 2013; Pavlenko, 2006).

One may think that the inverse relationship between population size and linguistic similarity observed in the present study may also reflect linguistic reflexes of modern African migrations to big cities. A large proportion of residents of bigger cities are usually migrants from small settlements. This may lead to linguistic similarity between small settlements and big cities which further results in a significant linguistic differences between big cities and medium-sized cities. However, this is

less likely to be the case in the Gurage areas. For instance, Silt'e, Inor and Mesqan are areas with the highest population size, but only small portion of the societies live in the cities; 3.3%, 2% and 39% respectively. The majority of these urban residents are speakers of Amharic. Hence, modern migration seems to have a marginal impact in the Gurage area. Probably migration in the past may have some contributions (see §2.2.2). Past migration was not just from small settlements to towns; it was also from village to towns and from towns to towns.

In general, it seems that Trudgill's Gravity model fails to take into account these factors. The re-consideration of such factors could better explain the nature of the relationship between linguistic distance and the population size. Assuming that the degree of influence between two centers is dependent on some of the quoted factors and that the relationship between linguistic distance and geographical distance is quadratic (regardless of the claims that the relationship can also be sub-linear), the Gravity model can be re-formulated as follow. Given,  $p$  = the population size,  $d$  = the distance between language sites,  $s$  = the similarity between a given language varieties and  $\epsilon$  = extra-linguistic factors, the influence between center  $i$  and center  $j$  is,

$$I_{ij} = \frac{s \cdot (\epsilon \pm (P_i P_j))}{d^2}$$

The plus or minus sign indicates the presence or the absence of the extra-linguistic factors. This formulation means that the linguistic similarity between related languages is always inversely proportional to the square of the geographical distance between two sites, but directly proportional to the product of the sizes of the population in the two sites, and several external variables that cannot be numerically expressed.

### 8.3 Limitations of the Tree Model

In 7.2.1, it was reported that the classifications of the South Ethiosemitic languages discussed in Chapter six are influenced by the contact among the speakers. The results of Neighbor-net representation and the lexicostatistical skewing (see §7.1.2.1



& 7.1.2.2) have also clearly illustrated that the horizontal diffusion of features is one of the major factors underpinning the classifications of the South Ethiosemitic languages. What is more interesting is that the classifications reported in Chapter six are very similar to the classifications previously reported by historical linguists based on the traditional comparative historical linguistics (e.g. Demeke, 2001; Hetzron, 1972; Hetzron, 1977). This triggers serious theoretical concerns about the distinction between tree-based genealogical classifications and wave-based areal classification of languages. As the basic assumptions of the Tree model and Wave model are quite different, the classifications based on the two models should not be the same. The Wave model refers to the horizontal transmissions of features while the Tree model considers the horizontal diffusion as a separate phenomena which is irrelevant for the reconstruction of genealogical relationships among languages (see §4.1). I would argue in this section that this similarity probably emerges from the flaws in the tree model, and that the classifications of Ethiosemitic languages previously reported by historical linguists may not so ‘genealogical’ as some scholars thought.

The cladistics (tree-based) representation of language classification presupposes an abrupt division of language community as the main cause of the emergence of a new language (François, 2015; Kalyan & François, 2018; Kalyan et al., 2018). Nevertheless, in reality, it is not always the case that speakers of a language split in such a manner that they lose contact once and forever. In other words, as also François (2015) observed, the Tree model fails to capture the very common situation in which linguistic diversification results from fragmentation of languages into networks of dialects which remain in contact with each other for an extended period of time. In the Tree model, each node represents a specific social community that evolved separately from another node which is the result of the actual split of a community due to various factors such as migration, blockage of communication or intrusion of other languages. This means that a given language may belong to only one higher-level group at a time; subgroups are mutually exclusive and never intersect. This principle of separate development is central to the whole arguments of sub-grouping studies under the cladistic approach (François, 2015). Historical linguists take this approach for granted and strive their best to exclude contact-induced

features in the genealogical classification of languages. Some others advise, at least, a separate treatment of the vertical and horizontal transmissions of features, but without providing a concrete way of distinguishing the two types of transmission.

A major shortcoming of the Tree model that emanates from the assumption of an abrupt social split is its inability to explain the relationship between innovations in two or more sister languages. For instance, as argued by François (2015), if there are shared innovations between sister languages  $m$  and  $n$ , the innovations are shared only by these two languages according to the Tree model. It fails to provide a concrete means of determining whether the shared innovations between the sister languages are the shared heritage of the remote ancestral language or not. The innovations between  $m$  and  $n$  can also be the outcomes of drift or parallel innovations. Alternatively, they can be transmitted via language contact.

The controversy of Wave vs. Tree model existed for a century and a half and the most common trend has been keeping the two approaches separate, i.e., employing the assumption of Tree model for genealogical classifications and the assumption of the Wave model for areal classification. However, making the distinction between the two has remained a challenge, especially in areas where there is a long history of language contact. Realizing these longstanding challenges, recently some linguists shifted their attentions towards adopting Wave-based approach for the genealogical classification of languages. One of the significant advancements in this regard is the emergence of Historical Glottometry (François, 2015; Kalyan & François, 2018; Kalyan et al., 2018). Historical Glottometry identifies sub-grouping in a <sup>19</sup>linkage situation and assesses their relative strength based on the distributions of innovations among languages. The model acknowledges the role of linguistic convergence and diffusion in the historical process of language diversification. It also recognizes the fact that languages usually consists networks of dialects, and language properties spread in space following complex patterns.

Historical Glottometry considers idiolects as the adequate unit of observation, not languages. It assumes that linguistic innovations first emerge in the speech of

---

<sup>19</sup>Linkage is network of dilaects which remained in contact with each other for an extened period of time

certain individuals, in the form of new ways of speaking. Then they spread among several speakers. After a period of competition with the previous norm, the innovation may become gradually dominant, and norm of the whole social group. From that point onward, the linguistic feature will be transmitted to the next generations. This language-internal diffusion of innovations gives rise to the genealogical relations among languages. This approach mainly focuses on language internal diffusion (diffusion among dialects of a language), not on the diffusion between separate languages. However, it does not consider contact between separate languages a threat to the genealogical classification. In short, this approach combines the historical comparative approach and the assumptions of the Wave model to determine the genealogical link among related languages.

Such a wave-based approach towards the genealogical classification of languages should be warmly welcomed since previous studies show that not all classifications of languages are effectively explained by the Tree model. For instance, the networks of Italian, Dutch and Arabic dialects could never be modeled by any tree (François, 2015, p.170). This casts also doubts about the accuracy of the Tree model when it is applied in the context of South Ethiosemitic languages. As discussed in 2.2.2, the Gurage area is a very diversified language area. Often each dialect consists of several sub-dialects. Given such a diversity and the long history of contact among the languages, the language situation deviates from the assumption of the Tree model in many ways. In such a small area with centuries of intense contact and borrowing among the languages, contact-free abrupt social split is simply inconceivable. In short, the linguistic and social reality in the Gurage contradicts the basic assumptions of the Tree model. Given these, it is also unlikely that previous classification proposals of the Ethiosemitic languages by historical linguists reflect the genealogical relationship among the languages, contrary to the predictions of historical linguists. The complex social and linguistic dynamics in the Gurage area requires a new perspective. The application of the approaches of Historical Glottometry on the South Ethiosemitic languages may open a new chapter towards dealing with the linguistic complexity of the area. It may also help to address the inconsistencies in the previous classifications of the languages which have been preserved probably due to the

misconception of the Tree model. I believe that future studies that employ the notion of Historical Glottometry will come up with a fresh idea about the genealogical classification of the South Ethiosemitic languages.

## 8.4 The Validity of Non-linguists' Knowledge

As discussed in 4.2.4, one of the debates in dialectology is whether the perceptual judgment of non-linguists can be used as a valid tool to determine the distance among related languages. Dialectologists are often divided on this idea. Some believe that native speakers of a language have an intuitive knowledge which can be used as a means of measuring linguistic distance while others express their doubt about considering such intuitive knowledge as reliable data for measuring the distance among related languages (see §4.2.4). Perceptual dialectologists tend to believe that perceptual distance which is measured based on authentic language inputs is more valid than perceptual distance which is measured without using language input, for example, by asking the non-native speakers in which close area a similar language is spoken. The present study employed recorded audio stimuli to determine the perceptual distance among the languages (see §6.3). While the classification of the languages based on the perceptual distance is largely similar to the classifications based on the two structural distances (phonetic and lexical), there are also differences. For example, the grouping of Silt'e with Kistane was attested only in the classification based on the perceptual distance. This was probably due to non-linguistic factors such as the attitude of the speakers. It was illustrated in 6.4 that there is very low consistency within the perceptual distance matrix. Besides, in Chapter seven, very strong correlation was reported between the perceptual distance and language attitude. These results clearly show that there are good reasons to be skeptical about the reliability of the perception-based linguistic distance measures.

The reliability of the perceptual distance may depend on several other factors. For instance, the higher the number of languages to be measured, the more the perceptual judgment may become inaccurate. This is probably because a judgment of similarity among many languages may increase memory load. The authenticity of

the test materials such as the quality of the recordings, loud speakers and headphones can also be a factor. This is probably the reason why in many of the previous studies (including the present study) the native speakers do not usually score 100% on the judgment of their own languages. Individual awareness of local varieties may also be hindered by several non-linguistic factors such as the degree of interaction among speakers of the varieties, geographical proximity among the speakers, degree of cultural and political cohesion among the speakers, landscape and availability of transportation that connects the communities.

In general, the accuracy of the perception-based distance measure is largely dependent on the type of perceptual test, the number of languages to be measured, the sociolinguistic and political situation of the language area and many other factors. This indicates that whenever the perceptual approach is opted to determine the distance among related languages, it must be employed with a high degree of caution. Perceptual approaches towards measuring linguistic distance shall be considered as an option only in a situation where employing the structural and functional approaches is not possible. As the perception of the speakers is largely associated with the political, cultural, social and economic situations of the speakers of the languages, the use of the perceptual approach to measure the distance among related languages should take into account all these factors.

## 8.5 The Influence of Language Attitude

The results reported in 7.2.3 show that there is a direct relationship between linguistic similarity and attitude of the speakers. In other words, the speakers are more positive about the languages that are structurally or functionally close to one's own language. The results do not reflect the speakers' negative attitude towards each others language as a result of political or social hostility which was reported in previous studies (Abu-Rabia, 1996; Abu-Rabia, 1998; Golubović & Sokolić, 2013). If this was the case, there could have been an inverse relationship between the linguistic similarity and the attitude of the speakers. Negative attitude of the speakers towards languages that are not familiar was interpreted as a reflection of the speakers'

conservative outlook towards a new language. The inverse correlation between the speakers' attitude and linguistic similarity is not always true though. This issue was illustrated in 7.2.3 using the data from Kistane and Silt'e speakers. The speakers of the two languages have a strong positive attitude towards each others' languages, but the languages are very dissimilar. This could be due to others extra-linguistic factors such as geographical adjacency.

It was also illustrated that the speakers' attitude correlates more strongly with the perceptual distance than with the functional and structural distances. This is a clear indication that the perceptual distance is primarily affected by language attitude compared to other dimensions of distance measures. In other words, the speakers' functional abilities associated with the linguistic distance are not strongly linked to the attitude of the speakers whereas the intuitive judgment of the speakers about the similarity between the language varieties is. This is probably because the attitude of speakers towards a language is usually associated not to linguistic factors, but to non-linguistic variables. These factors may not be necessarily negative; there may be bias because of positive social, political and economic factors.

The influence of language attitude can be minimal in the contexts where there is no past history of war among the speakers of the community, absence of historical hegemony of a dominant language, political and cultural antagonism among the speakers of the languages, and economic and social marginalization of a linguistic community (see §4.2.4). The opposite is true in the communities that have history of war and political hostility (see Abu-Rabia, 1996; Abu-Rabia, 1998; Golubović & Sokolić, 2013). The study by Golubović & Sokolić (2013) on Serbian-Croatian speakers indicates that animosity, negative political narratives and strong feeling of linguistic nationalism have a potential to reverse the widely reported positive attitude towards related languages. Feleke (2017) also reported the impact of the hostile political situation between Ethiopia and Eritrea on the attitude of the speakers. With regard to the Gurage varieties, though there were several events of war and conflicts in the history of the Gurage community (see §2.2.2), they are just events of remote past. For more than a century the linguistic landscape has not been as hostile as the Balkan areas. Therefore, it seems that the remote past con-

flicts among the Gurage community has a marginal influence on the attitude of the current generation of Gurage speakers.

## 8.6 The Combined Approach

It was illustrated in 1.1 and 4.3 that measuring linguistic distance is difficult for several reasons, but mainly because languages differ in multiple dimensions - phonetic, lexical, morphological and syntactic, and measuring the distance among related languages just from one perspective is not adequate. In order to overcome this one-dimensional nature of linguistic distance, the present study employed a combined approach - an approach which combines the structural, functional and perceptual dimensions.

The combined approach is essential to minimize the inherent limitations of each method of measuring linguistic distance that otherwise always remain a challenge. It is not possible to dissect, for instance, intelligibility from the non-linguistic factors - cognitive, psychological, cultural and so on. There is also no precise means of dealing with the asymmetry of intelligibility which was reported in 6.6 and in many other previous studies. Some scholars ( e.g., Tang & van Heuven, 2009) used the average of the two dimensions of intelligibility scores (the mean of the upper and the lower half of the intelligibility matrix) to deal with the problem of asymmetry. Nonetheless, since in principle, the mutual intelligibility is a bi-directional concept, the interpretation of such a mean can lead to conceptual problems.

In addition to the inherent problems in the intelligibility itself, preparing intelligibility tests that fairly consider the linguistic, cultural and social diversity of the speakers is extremely challenging (see §4.2.3). There are also practical problems in the administration of intelligibility tests; they are time and resource consuming. In this regard, the most serious problem that has received only a marginal attention in the previous studies is the *priming effect* which arises from repeating the same materials for the speakers of several languages. Repetition of the same materials can lead to familiarity to the test material by the test-takers which in turn may have a significant ramification on the intelligibility score. This could probably be

the case in the studies previously conducted on South Ethiosematic languages by Ahland (2003) and Menuta (2015). There are also challenges related to using the perceptual perspective of measuring linguistic distance. It was indicated in 6.3 and 8.4 that the perceptual perspective of measuring linguistic distance is very intuitive and is indeed influenced by the attitude of the speakers. Besides, it is the least reliable linguistic measure. This means that relying just on the perceptual measure can have a negative consequence (Schüppert & Gooskens, 2011; Van Bezooijen & Gooskens, 2007).

For various reasons such inherent limitations cannot be entirely avoided; they can only be reduced with careful material selection and administration. As indicated above, another means of minimizing these limitations is combining different aspects of distance measures. The combined approach provides a possibility to examine the distance from several dimensions and to generalize the distance among the languages based on multiple and adequate sets of data. Such an approach also creates a space for further inspection of the distance matrix and the detection of noises in the data. In this regard, the combined approach proposed in the present study paves a way for rigorous future investigations. Future studies may hopefully come up with a more systematic and improved ways of combining different distance measures.



---

## References

- Aalen, L. (2000). Ethiopia's paradox: Constitutional devolution and centralised party rule. In *Paper presented on the 14th international conference on Ethiopian studies*. November 2000, Addis Ababa University.
- Abu-Rabia, S. (1996). Attitudes and cultural background and their relationship to reading comprehension in a second language: A comparison of three different social contexts. *International Journal of Applied Linguistics*, 6(1), 81–105.
- Abu-Rabia, S. (1998). The influence of the Israel-Arab conflict on Israeli-Jewish students learning Arabic as a third language. *Language Culture and Curriculum*, 11(2), 154–164.
- Agee, J. (2018). *A Glottometric subgrouping of the Early Germanic languages* (Unpublished doctoral dissertation). San Jose State University.
- Ahland, M. (2003). Interlectal intelligibility between Gurage speech varieties. In *North American Conference on Afroasiatic linguistics* (pp. 45–71). April 4-6, Nashville.
- Ahland, M. (2004). *Language death in Mesmes: A sociolinguistic and historical-comparative examination of a disappearing language* (Unpublished doctoral dissertation). University of Texas at Arlington.
- Alewijnse, B., Nerbonne, J., Van der Veen, L., & Manni, F. (2007). A computational analysis of Gabon varieties. In *Proceedings of the ranlp workshop on computational phonology* (pp. 3–12). Recent Advances in Natural Language Phonology conference, Borovetz.

- Alexei, K., George, S., Anna, D., & Vasily, C. (2010). The Swadesh wordlist: An attempt at semantic specification. *Journal of Language Relationship*, 4, 46–89.
- Amberber, M., Baker, B., & Harvey, M. (2010). *Complex predicates: Cross-linguistic perspectives on event structure*. Cambridge: Cambridge University Press.
- Amha, A. (2010). Compound verbs and ideophones in Wolaitta revisited. In A. Mengistu, B. Brett, & H. Mark (Eds.), *Complex predicates: Cross-linguistic perspectives on event structure* (pp. 259–290). Cambridge/New York: Cambridge University Press.
- Anthony, L. (2005). Antconc: A learner and classroom friendly, multi-platform corpus analysis toolkit. In *Proceedings of iwlel* (pp. 7–13). Tokyo: Waseda University.
- Appleyard, D. (2015). Ethiopian Semitic and Cushitic. Ancient contact features in Ge'ez and Amharic. In M. Butts (Ed.), *Studies in Semitic languages and linguistics* (pp. 16–32). Leiden: Brill.
- April, M., Paul, H., Robert, M., & Natalia, S. (2005). Swadesh sublists and the benefits of borrowing: An Andean case study. *Transaction of the Philological Society*, 103(2), 147–170.
- Aurrekoetxea, G., Fernandez-Aguirre, K., Rubio, J., Ruiz, B., & Sánchez, J. (2013). 'diatech': A new tool for dialectology. *Literary and linguistic computing*, 28(1), 23–30.
- Axelsen, J. B., & Manrubia, S. (2014). River density and landscape roughness are universal determinants of linguistic diversity. *Proceedings of the Royal Society: Biological Sciences*, 281(1784), 1–8.
- Bailey, G. (1973). *Variation and linguistic theory*. Washington, DC: Georgetown University Press.
- Bailey, G., Wikle, T., Tillery, J., & Sand, L. (1993). Some patterns of linguistic diffusion. *Language variation and change*, 5(3), 359–390.

- Baker, J., & Pedersen, P. (1992). The rural-urban interface in Africa: Expansion and adaptation. Nordiska Afrikainstitutet. Denmark: Copenhagen.
- Banti, G. (2010). Remarks on the typology of converbs and their functional equivalents in East Cushitic. In A. Amaha, C. Rapold, & S. Zaugg-Coretti (Eds.), *Converbs, medial verbs, clause chaining and related issues* (pp. 31–80). Köln: Verlag.
- Beijering, K., Gooskens, C., & Heeringa, W. (2008). Predicting intelligibility and perceived linguistic distance by means of the Levenshtein algorithm. *Linguistics in the Netherlands*, 15, 13–24.
- Beijering, K., & Heeringa, W. (2007). *The role of phonetic and lexical distances in the intelligibility and perception of Scandinavian language varieties for speakers of standard Danish* (Unpublished doctoral dissertation). University of Groningen.
- Bender, M. (1966). Notes on lexical correlations in some Ethiopian languages. *Journal of Ethiopian Studies*, 4(1), 5–16.
- Bender, M. (1971). The languages of Ethiopia: A new lexicostatistic classification and some problems of diffusion. *Anthropological Linguistics*, 13(5), 165–288.
- Bender, M. (2003). Northeast Africa: A case study in genetic and areal linguistics. *Annual Publications in African Linguistics*, 1, 17–22.
- Bender, M., & Cooper. (1976). Language in Ethiopia: Implications of a survey for sociolinguistic theory and method. *Language in Society*, 1(2), 215–233.
- Bender, M., Cooper, R. L., & Ferguson, C. (1972). Language in Ethiopia: Implications of a survey for sociolinguistic theory and method. *Language in Society*, 1(2), 215–233.
- Bisang, W. (2006). Linguistic areas, language contact and typology: Some implications from the case of Ethiopia as a linguistic area. In M. Yaron (Ed.), *Linguistic areas* (pp. 75–98). Hampshire: Palgrave Macmillan.

- Bleses, D., Vach, W., Slott, M., Wehberg, S., Thomsen, P., Madsen, T. O., & Basbøll, H. (2008). Early vocabulary development in Danish and other languages: A CDI-based comparison. *Journal of child language*, *35*(3), 619–650.
- Bloem, J., Wieling, M., & Nerbonne, J. (2016). Automatically identifying characteristic features of non-native English accents. In M. Cote, J. Nerbonne, & K. Remco (Eds.), *The future of dialects* (pp. 155–173). Berlin: Language Science Press.
- Boberg, C. (2000). Geolinguistic diffusion and the US–Canada border. *Language variation and change*, *12*(1), 1–24.
- Boughton, Z. (2006). When perception isn't reality: Accent identification and perceptual dialectology in French. *Journal of French Language Studies*, *16*(3), 277–304.
- Bowern, C. (2012). The riddle of Tasmanian languages. *Proceedings of the Royal Society: Biological Sciences*, *279*(1747), 4590–4595.
- Bradlow, A., Clopper, C., Smiljanic, R., & Walter, M. A. (2010). A perceptual phonetic similarity space for languages: Evidence from five native language listener groups. *Speech Communication*, *52*(11-12), 930–942.
- Britain, D. (2017). Dialect contact and new dialect formation. In C. Bober, J. Nerbonne, & D. Watt (Eds.), *The hand book of dialectology* (pp. 143–158). Oxford: Wiley Blackwell.
- Bucholtz, M., Bermudez, N., Fung, V., Edwards, L., & Vargas, R. (2007). Hella Nor Cal or totally So Cal? the perceptual dialectology of California. *Journal of English Linguistics*, *35*(4), 325–352.
- Callary, R. (1975). Phonological change and the development of an urban dialect in Illinois. *Language in Society*, *4*(2), 155–169.
- Casad, E. (1987). *Dialect intelligibility test* (D. Evine, Ed.). Summer Institute of Linguistics. Dallas, Texas.

- Cavalli-Sforza, L., Menozzi, P., Cavalli-Sforza, L., Piazza, A., & Cavalli-Sforza, L. (1994). *The history and geography of human genes*. Princeton/New Jersey, Princeton University Press.
- Charles, B., Nerbonne, J., & Watt, D. (2018). Introduction. In C. Boberg, J. Nerbonne, & D. Watt (Eds.), *Handbook of Dialectology* (pp. 1–13). Hoboken, NJ: Wiley-Blackwell.
- Cheng, C. (1997). Measuring relationship among dialects: DOC and related resources. *International Journal of Computational Linguistics & Chinese Language Processing*, 2(1), 41–72.
- Chiswick, B., & Miller, P. W. (2005). Linguistic distance: A quantitative measure of the distance between English and other languages. *Journal of Multilingual and Multicultural Development*, 26(1), 1–11.
- Cohen, D. (1961). Le vocabulaire de base sémitique et le classement des dialectes du sud. *Semitica*, 11, 55–84.
- Crass, J., & Meyer, R. (2001). The Qabena and the Wolane: Two peoples of the Gurage region and their respective histories according to their own oral traditions. In *Annales d'éthiopie* (Vol. 17, pp. 175–182). Editions de la Table Ronde.
- Crass, J., & Meyer, R. (2009). Introduction. In J. Crass & R. Meyer (Eds.), *Language contact and language change in Ethiopia* (pp. 1–16). Cologne: Deutsche Nationalbibliothek.
- Demeke, G. (2001). The Ethio-Semitic languages (Re-examining the classification). *Journal of Ethiopian Studies*, 34(2), 57–93.
- Dinberu, A., Mengistu, H., Tesfaye, M., Tezerach, B., Girma, T., & Mitiku, T. (1995). *Gogot: YeGurage Bihereseb Tarik, Bahilina Qwanqwa (transliteration of Amharic)*. Addis Ababa: Artistic Printers.
- Drewes, A. (1996). The story of Joseph in Silt'i Gurage. In G. Hudson (Ed.), *Essays on Gurage language and culture* (pp. 69–92). Wiesbaden: Harrassowitz.

- Ehret, C. (1988). Social transformations in the early history of the Horn of Africa: Linguistic clues to developments of the period 500 BC to AD 500. In *Proceedings of eighth international conference of Ethiopian studies* (pp. 639–651). Institute of Ethiopian Studies Addis Ababa, Ethiopia.
- Eifert, B., Miguel, E., & Posner, D. N. (2010). Political competition and ethnic identification in Africa. *American Journal of Political Science*, *54*(2), 494–510.
- Embleton, S. M. (1982). *Incorporating borrowing rates in lexicostatistical tree reconstruction* (Unpublished doctoral dissertation). University of Tronto.
- ENSA. (2007). *Ethiopian statistical report of the 2007 population and housing census: Population size by age and sex, South Nation Nationalities*. Addis Ababa: Federal Democratic Republic of Ethiopia, Population Census Commission.
- Faber, A. (1997). Genetic subgrouping of the Semitic languages. *The Semitic Languages*, *15*, 3–15.
- Feleke, T. L. (2017). The similarity and mutual intelligibility between Amharic and Tigrigna varieties. In *Proceedings of the fourth workshop on nlp for similar languages, varieties and dialects (vardial)* (pp. 47–54). April, 2017, Spain Valencia.
- Felhnan, J. (1996). Lines on the classification of Ethiopian-Semitic. *Studies in African Linguistics*, *25*(2), 205–206.
- Ferguson, C. (1976). The Ethiopian language area. Language in Ethiopia. In M. Bender, J. Bown, R. Cooper, & C. Ferguson (Eds.), *Language in Ethiopia* (pp. 63–76). Oxford: Oxford University Press.
- Fleming, H. (1968). Ethiopic language history: Testing linguistic hypotheses in an archaeological and documentary context. *Ethnohistory*, *15*(4), 353–388.
- Fossen, K. (2018). *The intelligibility of Norwegian-accented English* (Master's Thesis). University of Oslo, Norway.

- François, A. (2015). Trees, waves and linkages: Models of language diversification. In B. Charless & E. Brown (Eds.), *The Routledge handbook of historical linguistics* (pp. 161–189). London: Routledge.
- Geisler, H., & List, J.-M. (2013). Do languages grow on trees? The tree metaphor in the history of linguistics. In H. Fangerau, H. Geisler, T. Halling, & W. Martin (Eds.), *Classification and evolution in biology, linguistics and the history of science* (pp. 111–124). Stuttgart: Ranz Steiner Verlag,.
- Goebel, H. (1982). Ansätze zu einer computativen dialektometrie. *Ein Handbuch zur deutschen und allgemeinen Dialektforschung. Handbücher zur Sprach-und Kommunikationswissenschaft, 1*, 778–792.
- Goebel, H. (2010). Dialectometry and quantitative mapping. In A. Lameli, R. Kehrein, & S. Rabanus (Eds.), *Language and space. An international handbook of linguistic variation* (Vol. 2, pp. 433–457). Berlin, New York: De Gruyter Mouton.
- Goeman, A. (1999). Dialects and subjective judgement of speakers. In L. Daniel & D. Preston (Eds.), *Handbook of perceptual dialectology* (Vol. 2, pp. 135–136). Amsterdam: John Benjamins.
- Goldenberg, G. (1977). The Semitic languages of Ethiopia and their classification. *Bulletin of the School of Oriental and African Studies, 40*(3), 461–507.
- Golubović, J., & Sokolić, N. (2013). Their language sounds aggressive: A matched guise study with Serbian and Croatian. In C. Gooskens & R. Van Bezooijen (Eds.), *Phonetics in Europe. Perception and Production* (pp. 35–57). Frankfurt: Peter Lang.
- Goosens, J. (1977). *Inleiding tot de nederlandse dialectologie*. Groningen: Wolters-Noordhoff.
- Gooskens, C. (1997). On the role of prosodic and verbal information in the perception of Dutch and English varieties (PhD Dissertation). University of Nijmegen.

- Gooskens, C. (2005a). How well can Norwegians identify their dialects? *Nordic Journal of Linguistics*, 28(1), 37–60.
- Gooskens, C. (2005b). Traveling time as a predictor of linguistic distance. *Dialectologia et Geolinguistica*, 13, 38–62.
- Gooskens, C. (2007). The contribution of linguistic factors to the intelligibility of closely related languages. *Journal of Multilingual and multicultural development*, 28(6), 445–467.
- Gooskens, C. (2013). Experimental methods of measuring intelligibility of closely related language varieties. In B. Robert, C. Richard, & L. Ceil (Eds.), *The Oxford Handbook of Sociolinguistics* (pp. 195–213). Oxford: Oxford University Press.
- Gooskens, C. (2018). Dialect intelligibility. In C. Boberg, J. Nerbonne, & D. Watt (Eds.), *Handbook of Dialectology* (pp. 204–218). Hoboken, NJ: Wiley-Blackwell.
- Gooskens, C., & Heeringa, W. (2004). Perceptive evaluation of Levenshtein dialect distance measurements using Norwegian dialect data. *Language variation and change*, 16(3), 189–207.
- Gooskens, C., Heeringa, W., & Beijering, K. (2008). Phonetic and lexical predictors of intelligibility. *International Journal of Humanities and Arts Computing*, 2(1-2), 63–81.
- Gooskens, C., & Heuven, V. (2018). How well can intelligibility of closely related languages in Europe be predicted by linguistic and non-linguistic variables? *Linguistic Approaches to Bilingualism*, 4(8), 1-29.
- Gooskens, C., & Van Bezooijen, R. (2006). Mutual comprehensibility of written Afrikaans and Dutch: Symmetrical or asymmetrical? *Literary and Linguistic Computing*, 21(4), 543–557.
- Gooskens, C., Van Heuven, V. J., Van Bezooijen, R., & Pacilly, J. J. (2010). Is spoken Danish less intelligible than Swedish? *Speech Communication*, 52(11-12), 1–33.



- Gould, P., & White, R. (1974). *Mental maps*. Newyork: Penguin Books Inc.
- Gragg, G. (1997). Old South Arabian Phonology. In A. Keye (Ed.), *Phonology of Asia and Africa* (Vol. 1, pp. 163–165). Indiana: Winona Lake.
- Grimes, J. (1990). Interpreting sample variation in intelligibility tests. In G. Williams (Ed.), *Synchronic and diachronic approaches to linguistic variation and change* (pp. 138–146). Washington, DC: Georgetown University.
- Grimes, J. (1995). *Language survey reference guide*. Dalla: Summer Inst of Linguistics.
- Güldemann, T., & Fiedler, I. (2019). Niger-Congo “noun classes” conflate gender with deriflection. *Grammatical gender and linguistic complexity*, 1, 95–145.
- Gutt, E.-A. (1980). Intelligibility and interlingual comprehension among selected Gurage speech varieties. *Journal of Ethiopian Studies*, 14, 57–85.
- Hamed, M. B. (2005). Neighbour-nets portray the Chinese dialect continuum and the linguistic legacy of China’s academic history. *Proceedings of the Royal Society of London: Biological Sciences*, 272(1567), 1015–1022.
- Heeringa, W. (2004). *Measuring dialect pronunciation differences using Levenshtein distance* (Unpublished doctoral dissertation). University of Groningen.
- Heeringa, W., Kleiweg, P., Gooskens, C., & Nerbonne, J. (2006). Evaluation of string distance algorithms for dialectology. In *Proceedings of the workshop on linguistic distances* (pp. 51–62). Sydney, Association for Computational Linguistics.
- Heeringa, W., & Nerbonne, J. (2000). Change, convergence and divergence among Dutch and Frisian. In B. Piter, B. Piter, J. Lammert, & V. Jacob (Eds.), *Philologia frisica anno 1999. lêzingen fan it fyftjinfryskfilologekongres* (pp. 88–109). Ljouwert: Fryske Akademy.
- Heeringa, W., Nerbonne, J., van Bezooijen, R., & Spruit, M. R. (2011). Geografie en inwoneraantallen als verklarende factoren voor variatie in het Nederlandse dialectgebied. *Tijdschrift voor Nederlandse taal-en letterkunde*, 123(1), 70–82.

- Heeringa, W., Swarte, F., Schüppert, A., & Gooskens, C. (2014). Modeling intelligibility of written Germanic languages: Do we need to distinguish between orthographic stem and affix variation? *Journal of Germanic Linguistics*, 26(4), 361–394.
- Heeringa, W., Swarte, F., Schüppert, A., & Gooskens, C. (2018). Measuring syntactical variation in Germanic texts. *Digital Scholarship in the Humanities*, 33(2), 279–296.
- Heggarty, P. (2010). Beyond lexicostatistics: How to get more out of ‘word list’ comparisons. *Diachronica*, 27(2), 301–324.
- Heine, B. (1974). Historical linguistics and lexicostatistics in Africa. *Journal of African Linguistics*, 3(11), 7–20.
- Henry, W. (2001). *Doing development and being Gurage: The embeddedness of development in Sebat Bet Gurage identities* (Unpublished doctoral dissertation). Open University.
- Hetzron, R. (1968). Main verb-markers in Northern Gurage. *Africa*, 38(2), 156–172.
- Hetzron, R. (1972). *Ethiopian Semitic: Studies in classification* (No. 2). Manchester University Press.
- Hetzron, R. (1977). *The Gunnän-Gurage languages* (Vol. 12). Naples: Istituto Universitario Orientale.
- Hetzron, R., & Bender, M. (1976). The Ethiosemitic languages. In L. Bender (Ed.), *Languages in Ethiopia* (pp. 2–29). London: Oxford University Press.
- Hilton, N. H., Gooskens, C., & Schüppert, A. (2013). The influence of non-native morphosyntax on the intelligibility of a closely related language. *Lingua*, 137, 1–18.
- Hinnebusch, T. J. (1996). Skewing in lexicostatistic tables as an indicator of contact. *Round Table on Bantu Historical Linguistics, Université Lumière, 2, Lyon, France, 30 May–1 June 1996*.

- Horvath, B., & Horvath, R. (2001). A multilocality study of a sound change in progress: The case of /l/ vocalization in New Zealand and Australian English. *Language variation and change*, 13(1), 37–57.
- Hudson, G. (1977). Language classification and the Semitic prehistory of Ethiopia. *Folia Orientalia Krakow*, 18, 119–166.
- Hudson, G. (2000). Ethiopian Semitic overview. *Journal of Ethiopian Studies*, 33(2), 75–86.
- Hudson, G. (2013). *Northeast African Semitic: Lexical comparisons and analysis*. Wiesbaden: Harrassowitz.
- Hudson, G. (2018). Lexicostatistical evidence for Ethiosemitic, its subgroups, and borrowing. In S. Joseph (Ed.), *Afroasiatic: Current issues in linguistic theory* (pp. 41–61). Madison: University of Wisconsin.
- Huson, D. H., & Bryant, D. (2010). *User manual for SplitsTree4 v4. 13.1*. Unpublished.
- Hutchinson, W. K. (2005). "linguistic distance" as a determinant of bilateral trade. *Southern Economic Journal*, 72(1), 1–15.
- Kalyan, S., & François, A. (n.d.). When the waves meet the trees: A response to Jacques and List. *Journal of Historical Linguistics*, 9(1), 167–179.
- Kalyan, S., François, A., et al. (2018). Freeing the Comparative Method from the Tree Model: A framework for Historical Glottometry. *Senri Ethnological Studies*, 98, 59–89.
- Kessler, B. (1995). Computational dialectology in Irish Gaelic. In *Proceedings of the seventh conference on European Chapter of Association for Computational Linguistics* (pp. 60–66). Burlington: Morgan Kaufmann.
- Kirk, P. (1970). Dialect intelligibility testing: The Mazatec study. *International Journal of American Linguistics*, 36(3), 205–211.

- Kitchen, A., Ehret, C., Assefa, S., & Mulligan, C. J. (2009). Bayesian phylogenetic analysis of Semitic languages identifies an Early Bronze Age origin of Semitic in the Near East. *Proceedings of the Royal Society of London: Biological Sciences*, 276, 2703–2710.
- Kogan, L. (2005). Common origin of Ethiopian Semitic: The lexical dimension. *Scrinium*, 1(1), 367–396.
- Kremer, L. (1999). The Netherlands-German national border as a subjective dialect boundary. In D. Preston (Ed.), *Handbook of perceptual dialectology* (Vol. 1, pp. 31–36). Amsterdam: Benjamins.
- Kretzschmar, W. A. (1992). Isoglosses and predictive modeling. *American Speech*, 67(3), 227–249.
- Kürschner, S., Gooskens, C., & van Bezooijen, R. (2008). Linguistic determinants of the intelligibility of Swedish words among Danes. *International Journal of Humanities and Arts Computing*, 2(1-2), 83–100.
- Labov, W. (2001). Principles of language change: Social factors. *Malden, MA: Blackwell*.
- Labov, W. (2011). *Principles of linguistic change: Cognitive and cultural factors* (Vol. 3). West Sussex: John Wiley & Sons.
- Leinonen, T., Çoltekin, Ç., & Nerbonne, J. (2016). Using Gabmap. *Lingua*, 178, 71–83.
- Leslau, W. (1945). The influence of Cushitic on the Semitic languages of Ethiopia, a problem of substratum. *Word*, 1(1), 59–82.
- Leslau, W. (1951). Archaic features in South Ethiopic. *Journal of the American Oriental Society*, 71(4), 212–230.
- Leslau, W. (1952). The influence of Sidamo on the Ethiopic languages of Gurage. *Language*, 28(1), 63–81.

- Leslau, W. (1956). The Arabic loanwords in Gurage (Southern Ethiopia). *Arabica*, 3(3), 266–284.
- Leslau, W. (1967). Hypothesis on a Proto-Semitic marker of the imperfect in Gurage. *Journal of Near Eastern Studies*, 26(2), 121–125.
- Leslau, W. (1969). Toward a classification of the Gurage dialects. *Journal of Semitic Studies*, 14(1), 96–109.
- Leslau, W. (1979). *Etymological Dictionary of Gurage (Ethiopic): English-Gurage index* (Vol. 2). Otto Harrassowitz Verlag.
- Leslau, W. (1982). The jussive in the Gurage dialects of Muher and Masqan. *Zeitschrift der deutschen morgenländischen Gesellschaft*, 132(1), 85–97.
- Leslau, W. (1992). *Gurage studies: Collected articles*. Otto Harrassowitz Verlag.
- Leslau, W. (1996). Chaha and Ennemor: An analysis of two Gurage dialects. In G. Hudson (Ed.), *Essays on Gurage language and culture* (pp. 111–122). Wiesbaden: Harrassowitz.
- Levenshtein, V. (1966). Binary codes capable of correcting deletions, insertions, and reversals. In *Soviet physics doklady* (Vol. 10, pp. 707–710). Original in Russian in Dokl.Akad.Naur SSSSR 163, 4, 845-848, 1965.
- Leyew, Z., & Siebert, R. (2001). *Sociolinguistic survey report of the Argobba language of Ethiopia*. Addis Ababa: SIL International.
- Little, G. (1974). Syntactic evidence of language contact: Cushitic influence in Amharic. In S. Roger & B. Charles-James (Eds.), *Towards tomorrow's linguistics* (pp. 267–275). Washington DC: Georgetown University Press.
- Long, D. (1999a). Geographical perceptions of Japanese dialect regions. In D. Preston (Ed.), *Handbook of perceptual dialectology* (Vol. 1, pp. 177–198). Amsterdam: John Benjamins.

- Long, D. (1999b). Mapping nonlinguists' evaluations of Japanese language variation. In D. Preston (Ed.), *Handbook of perceptual dialectology* (Vol. 1, pp. 199–226). Amsterdam: John Benjamins.
- Marcos, H. M. (1973). Regional variations in Amharic. *Journal of Ethiopian Studies*, 11(2), 113–129.
- Mase, Y. (1999). On dialect consciousness dialect characteristics given by speakers. In D. Preston (Ed.), *Handbook of perceptual dialectology* (Vol. 1, pp. 101–113). Amsterdam: John Benjamins.
- Menuta, F. (2015). *Intergroup communication among Gurage: A Study in Intelligibility, Inter-lingual Comprehension and Accommodation*. Saarbrücken: Lambert Academic Publishing.
- Meyer, R. (2005a). Das zay. deskriptive grammatik einer ostguragesprache (äthiosemitisch). *Aethiopica*, 8, 293–294.
- Meyer, R. (2005b). Riddles as indicator of cultural and linguistic convergence in the Gurage region. *Scrinium*, 1(1), 174–196.
- Meyer, R. (2006). *Wolane: descriptive grammar of an East Gurage language (Ethiosemitic)*. Köln: Köppe Verlag.
- Meyer, R. (2011). Gurage. In S. Weninger (Ed.), *The Semitic languages: An international handbook* (pp. 1220–1257). Berlin: DeGruyter Mouton.
- Meyer, R. (2014). Finiteness in Gurage languages. In A. Amha, R. Meyer, & T. Yvonne (Eds.), *Explorations in Ethiopian linguistics: Complex predicates, finiteness and interrogativity* (pp. 225–258). Wiesbaden: Harrassowitz.
- Meyer, R. (2018). Language standardization efforts in Gurage. *Conference paper presented on Ethiopian Studies (October, 1–5)*. Mekelle: Mekelle University.
- Milroy, J., & Milroy, L. (2012). *Authority in language: Investigating standard English*. London: Routledge.

- Minett, J., & Wang, W. (2003). On detecting borrowing: Distance-based and character-based approaches. *Diachronica*, 20(2), 1–80.
- Montgomery, C. (2007). *Northern English dialects: A perceptual approach* (Unpublished doctoral dissertation). University of Sheffield.
- Moulton, W. G. (1960). The short vowel systems of Northern Switzerland: A study in structural dialectology. *Word*, 16(2), 155–182.
- Negesse, F. (2015). Classification of Oromo dialects: A computational approach. *International Journal of Computational Linguistics (IJCL)*, 6(1), 1–10.
- Nerbonne, J. (2010). Measuring the diffusion of linguistic change. *Philosophical Transactions of the Royal Society of London: Biological Sciences*, 365(1559), 3821–3828.
- Nerbonne, J., Colen, R., Gooskens, C., Kleiweg, P., & Leinonen, T. (2011). Gabmap—a web application for dialectology. *Dialectologia: Revista electrònica*, 2(2), 65–89.
- Nerbonne, J., & Heeringa, W. (2001). Computational comparison and classification of dialects. *Dialectologia et Geolinguistica*, 9, 69–83.
- Nerbonne, J., & Heeringa, W. (2010). Measuring dialect differences. In P. Auer & J. Schmedit (Eds.), *Language and space: An instructional handbook of linguistic variation* (Vol. I, pp. 550–592). De Gruyter Mouton: Berlin/New York.
- Nerbonne, J., Heeringa, W., & Kleiweg, P. (1999). Edit distance and dialect proximity. In D. Sankoff & P. Kruskal (Eds.), *Time warps, string edits and macromolecules: The theory and practice of sequence comparison* (pp. v–xv). CSLI: Stanford.
- Nerbonne, J., & Hinrichs, E. (2006). Linguistic distances. In *Proceedings of the workshop on linguistic distances* (pp. 1–6). USA: Stroudsburg.
- Nerbonne, J., & Kleiweg, P. (2003). Lexical distance in LAMSAS. *Computers and the Humanities*, 37(3), 339–357.

- Nerbonne, J., & Kretzschmar, W. (2003). Introducing computational techniques in dialectometry. *Computers and the Humanities*, 37(3), 245–255.
- Nerbonne, J., & Siedle, C. (2005). Dialektklassifikation auf der grundlage aggregierter ausspracheunterschiede. *Zeitschrift für Dialektologie und Linguistik*, 72(2), 129–147.
- Nerbonne, J., Van Gemert, I., & Heeringa, W. (2005). *A dialectometric view of linguistic “gravity”*. Groningen: University of Groningen.
- Nerbonne, J., & Wiersma, W. (2006). A measure of aggregate syntactic distance. In J. Nerbonne & I. Hinrichs (Eds.), *Linguistic distances. workshop at the joint conference of international committee on computational linguistics and the association for computational linguistics* (pp. 82–90). Sidney: Australia.
- Nida, W. (1996). *Gurage urban migration and the dynamics of cultural life in the village* (G. Hudson, Ed.). Wiesbaden: Harrassowitz.
- Nida, W. (2000). Fanonet: Ethnohistorical notes on the Gurage urban migration in Ethiopia. *Ufahamu: A Journal of African Studies*, 28(34-73).
- Niedzielski, N. (1999). The effect of social information on the perception of sociolinguistic variables. *Journal of language and social psychology*, 18(1), 62–85.
- Ousman, S. (2015). Descriptive grammer of Mesqan (PhD Thesis). Addis Ababa: Addis Ababa University.
- Palmer, F. (1958). Comparative statement and Ethiopian Semitic. *Transactions of the Philological Society*, 57(1), 119–143.
- Palmer, F. (1978). Hetzron Robert: The Gunnän-Gurage languages.(Ricerche: Studi diSemitistica e del Vicino Oriente Antico, XII.) X, 264 pp. Napoli: Istituto Orientale di Napoli, 1977. *Bulletin of the School of Oriental and African Studies*, 41(3), 584–585.
- Pavlenko, A. (2006). Russian as a lingua franca. *Annual Review of Applied Linguistics*, 26, 78–99.



- Pearce, M. (2009). A perceptual dialect map of North East England. *Journal of English Linguistics*, 37(2), 162–192.
- Petroni, F., & Serva, M. (2010). Measures of lexical distance between languages. *Physica: Statistical Mechanics and its Applications*, 389(11), 2280–2283.
- Plichta, B., & Preston, D. (2005). The /ay/s have it the perception of /ay/ as a north-south stereotype in United States English. *Acta Linguistica Hafniensia*, 37(1), 107–130.
- Preston, D. (1993). Folk dialectology. In D. Preston (Ed.), *American dialect research* (pp. 333–337). Amsterdam: Benjamins.
- Preston, D. (1996). Where the worst English is spoken. In E. Schneider (Ed.), *Focus on USA* (pp. 297–361). Amsterdam: John Benjamins.
- Preston, D. (1999a). The discussions surrounding the subjective boundaries of dialects. In D. Preston (Ed.), *Handbook of perceptual dialectology* (Vol. 1, pp. 115–130). Amsterdam: John Benjamins.
- Preston, D. (1999b). Introduction. In D. Preston (Ed.), *Handbook of perceptual dialectology* (Vol. 1, pp. xxv–xl). Amsterdam: John Benjamins.
- Preston, D. (2010). Mapping the geolinguistic space of the brain. In A. Lameli, R. Kehrein, & S. Rabanus (Eds.), *Language and space. An international handbook of linguistic variation: Language mapping* (Vol. 2, pp. 121–141). Berlin, New York: De Gruyter Mouton.
- Preston, D. (2018). Perceptual dialectology. In C. Boberg, J. Nerbonne, & D. Watt (Eds.), *Handbook of Dialectology* (pp. 173–207). Hoboken, NJ: Wiley-Blackwell.
- Prokić, J. (2007). Identifying linguistic structure in a quantitative analysis of dialect pronunciation. In *Proceedings of the 45th annual meeting of the acl: Student research workshop* (pp. 61–66). Prague: Association for Computational Linguistics.

- Prokić, J., Çöltekin, Ç., & Nerbonne, J. (2012). Detecting shibboleths. In *Proceedings of the EACL 2012 joint workshop of LINGVIS & UNCLH* (pp. 72–80). Stroudsburg: Association for Computational Linguistics.
- Prokić, J., & Nerbonne, J. (2008). Recognizing groups among dialects. *International Journal of Humanities and Arts Computing*, 2(1-2), 153–172.
- Prokić, J., Nerbonne, J., et al. (2013). Analyzing dialects biologically. In F. Heiner, G. Hans, H. Thoresten, & M. William (Eds.), *Classification and evolution in biology, linguistics and the history of science* (pp. 147–159). Stuttgart: Franz Steiner.
- Rabanus, S. (2018). Dialect Maps. In C. Boberg, J. Nerbonne, & D. Watt (Eds.), *Handbook of Dialectology* (pp. 348–367). Hoboken, NJ: Wiley-Blackwell.
- Renée, v. B., & Van den Berg, R. (1999). Word intelligibility of language varieties in the Netherlands and Flanders under minimal conditions. *Linguistics in the Netherlands*, 16(1), 1–12.
- Rose, S. (1996). Allomorphy and morphological categories in Muher. In G. Hudson (Ed.), *Essays on Gurage language and culture* (pp. 205–227). Wiesbaden: Harrassowitz.
- Rose, S. (2007). Chaha (Gurage) morphology. In S. Kaye (Ed.), *Morphologies of Asia and Africa* (pp. 403–427). Winona Lake: Eisenbrauns.
- Rubin, A. D. (2008). The subgrouping of the Semitic languages. *Language and Linguistics Compass*, 2(1), 79–102.
- Samia, Z. (2007). *Statistical report for Southern Nations, Nationalities and Peoples' Region*. Addis Ababa: UNFPA.
- Sampson, G. (2000). A proposal for improving the measurement of parse accuracy. *International Journal of Corpus Linguistics*, 5(1), 53–68.
- Sanders, N. (2007). Measuring syntactic difference in British English. In *Proceedings of the acl 2007 student research workshop* (pp. 1–7). Madison, Omni-press.

- Sankoff, D. (1972). Reconstructing the history and geography of an evolutionary tree. *The American Mathematical Monthly*, 79(6), 596–603.
- Scassa, T. (1994). Language standards, ethnicity and discrimination. *Canadian Ethnic Studies*, 26(3), 105–121.
- Schepens, J., Dijkstra, T., & Grootjen, F. (2012). Distributions of cognates in Europe as based on Levenshtein distance. *Bilingualism: Language and Cognition*, 15(1), 157–166.
- Schleicher, A. (1853). The first divisions of the Indo-european indigenous people[English translation]. *General Monthly Bulletin for Science and Literature[English Translation]*, 786-787.
- Schmidt, J. (1872). *Die verwantschaftsverhältnisse der indogermanischen sprachen*. Weimar, Germany.
- Schneider, E. W. (2003). The dynamics of New Englishes: From identity construction to dialect birth. *Language*, 79(2), 233–281.
- Schüppert, A., & Gooskens, C. (2011). Investigating the role of language attitudes for perception abilities using reaction time. *Dialectologia: Revista electrònica*, 2, 119–140.
- Schüppert, A., Hilton, N. H., & Gooskens, C. (2015). Swedish is beautiful, Danish is ugly? Investigating the link between language attitudes and spoken word recognition. *Linguistics*, 53(2), 375–403.
- Séguy, J. (1973). La dialectométrie dans l'Atlas linguistique de la Gascogne. *Revue de linguistique romane*, 37, 1–24.
- Séguy, J., & Ravier, X. (1973). *Atlas linguistique de la gascogne.* "Notice explicative" (Vol. 6). Paris: Centre National de la Recherche Scientifique.
- Serva, M., & Petroni, F. (2008). Indo-European languages tree by Levenshtein distance. *EPL (Europhysics Letters)*, 81(6), 1–8.

- Shack, W. (1966). *The Gurage: A people of the ensete culture*. UK: Routledge.
- Sibata, T. (1999). Consciousness of dialect boundaries. In D. Preston (Ed.), *Handbook of perceptual dialectology* (Vol. 1, pp. 39–69). Amsterdam: John Benjamins.
- Silvestri, M., & Tomezzoli, G. (2005). Linguistic computational analysis to measure the distances between ancient Venetic, Latin and Slovenian Languages. In *Proceedings of the third international topical conference, ancient settlers of europe* (pp. 77–85). Jutro, Ljubljana.
- Snoek, C. (2014). Review of Gabmap: Doing dialect analysis on the web. *Language documentation and conservation*, 8, 192–208.
- Spruit, M. R. (2005). Classifying Dutch dialects using a syntactic measure: The perceptual Daan and Block dialects map revisited. *Linguistics in the Netherlands*, 22(1), 179–190.
- Spruit, M. R. (2006). Measuring syntactic variation in Dutch dialects. *Literary and linguistic Computing*, 21(4), 493–506.
- Storch, A., & Dimmendaal, G. J. (2014). One size fits all. *Number constructions and semantics: Case studies from Africa, Amazonia, India, and Oceania*, 3(8), 1–32.
- Swadesh, M. (1950). Salish internal relationships. *International Journal of American Linguistics*, 16(4), 157–167.
- Swadesh, M. (1955). Towards greater accuracy in lexicostatistic dating. *International journal of American linguistics*, 21(2), 121–137.
- Tadesse, M. (2018). *Documentation & description of Amharic dialects* (Unpublished doctoral dissertation). Addis Ababa: Addis Ababa University.
- Tamasi, S. (2003). *Cognitive patterns of linguistic perceptions* (Unpublished doctoral dissertation). Georgia: University of Georgia.
- Tang, C., & van Heuven, V. (2009). Mutual intelligibility of Chinese dialects experimentally tested. *Lingua*, 119(5), 709–732.

- Tang, C., & van Heuven, V. J. (2007). Mutual intelligibility and similarity of Chinese dialects: Predicting judgments from objective measures. *Linguistics in the Netherlands*, 24(1), 223–234.
- Tang, C., & van Heuven, V. J. (2015). Mutual intelligibility of Chinese dialects: Predicting cross-dialect word intelligibility from lexical and phonological similarity. In *Proceedings of the 18th international congress of phonetic science* (pp. 1–5). UK, Glasgow.
- Tang, C., & Van Heuven, V. J. (2015). Predicting mutual intelligibility of Chinese dialects from multiple objective linguistic distance measures. *Linguistics*, 53(2), 285–312.
- Taye, A. (1954). *History of Ethiopian people [translated]*. Addis Ababa: Commercial Printing Press.
- Tesema, A. B., Ann, B., & Bo, T. (1993). *Useful trees and shrubs for Ethiopia: Identification, propagation and management for agricultural and pastoral communities* (T. Kelbesa, S. Demissew, & P. Maundu, Eds.). Nairobi: World Agroforestry Centre.
- Tilahun, W. (2015). *Documentation and description of traditional conflict resolution of Sebat Bet Gurage: The Case of Chaha Gurage* (Master's Thesis). Addis Ababa University, Addis Ababa.
- Tosco, M. (2000). Is there an "Ethiopian language area"? *Anthropological Linguistics*, 42(3), 329–365.
- Trudgill, P. (1974). Linguistic change and diffusion: Description and explanation in sociolinguistic dialect geography. *Language in society*, 3(2), 215–246.
- Trudgill, P. (1983). *On dialect: Social and geographic perspectives*. New York: New York University Press.
- Trudgill, P. (2004). *New-dialect formation: The inevitability of colonial Englishes*. Oxford: Oxford University Press.

- Ullendorff, E. (1955). *The Semitic languages of Ethiopia: A comparative phonology*. London: Taylor's (Foreign) Press.
- Van Bezooijen, R., & Gooskens, C. (2007). Interlingual text comprehension: Linguistic and extralinguistic determinants. In T. Jan & Z. Ludger (Eds.), *Receptive multilingualism, linguistic analyses, language policies and dialectic concepts* (pp. 249–263). Amsterdam: John Benjamins.
- Voegelin, C. F., & Harris, Z. S. (1951). Methods for determining intelligibility among dialects of natural languages. *Proceedings of the American Philosophical Society*, 95(3), 322–329.
- Voigt, R. (2009). North vs. South Ethiopian Semitic. In *Proceedings of the 16th international conference of Ethiopian studies* (Vol. 4, pp. 1375–1387). Addis Ababa: Addis Ababa University.
- Völlmin, S. (2017). *Towards a Grammar of Gumer - phonology and morphology of a Western Gurage variety* (Unpublished doctoral dissertation). University of Zurich.
- Völlmin, S., Amha, A., Rapold, C. J., & Zaugg-Coretti, S. (2007). *Converbs, medial verbs, clause chaining and related issues* (Vol. 19). Köln: Rüdiger Köppe.
- Wang, W. (2007). *English as a lingua franca: Mutual intelligibility of Chinese, Dutch and American speakers of English*. Utrecht: Netherlands Graduate School of Linguistics.
- Wang, W., & Minett, J. (2005). Vertical and horizontal transmission in language evolution. *Transactions of the Philological Society*, 103(2), 121–146.
- Watt, D. (2018). Applied Dialectology: Dialect Coaching, Dialect Reduction, and Forensic Phonetics. In C. Boberg, J. Nerbonne, & D. Watt (Eds.), *Handbook of Dialectology* (pp. 219–232). Hoboken, NJ: Wiley-Blackwell.
- Wetter, A. (2007). Converbs in Argoba. In S. Völlmin, A. Amha, C. Rapold, & S. Zaugg-Coretti (Eds.), *Converbs, medial verbs, clause chaining and related issues* (Vol. 19, pp. 99–113). Köln: Köppe Verlag.

- Wieling, M., Heeringa, W., & Nerbonne, J. (2007). An aggregate analysis of pronunciation in the Goeman-Taeldeman-Van Reenen-Project data. *Taal en Tongval*, 59(1), 84–116.
- Wieling, M., Margareta, E., & Nerbonne, J. (2012). Inducing measure of phonetic similarity from pronunciation variation. *Journal of Phonetics*, 5(8), 1–41.
- Wieling, M., & Nerbonne, J. (2011). Bipartite spectral graph partitioning for clustering dialect varieties and detecting their linguistic features. *Computer Speech & Language*, 25(3), 700–715.
- Wikle, T. (1997). The spatial diffusion of linguistic features in Oklahoma. In *Proceedings of the Oklahoma Academy of Science* (Vol. 77, pp. 1–15). Available at digital.library.okstate.edu/OAS/.
- Wolyie, J. (2008). The politics of language, power and pedagogy in Ethiopia: Addressing the past and present conditions of the Oromo language. *Australian Journal of Linguistics*, 28(1), 31–57.
- Worku, N. (2005). Ethno-historical survey. In S. Uhlig (Ed.), *Encyclopaedia Aethiopica* (Vol. 2, pp. 929–933). Harrassowitz: Verlag.
- Zaborski, A. (1991). Ethiopian language subareas. In P. Stanisław & R. Eugeniusz (Eds.), *Unwritten testimonies of the African past* (pp. 123–134). Warsaw: Wydawnictwa Uniwersytetu Warszawskiego.
- Zealelem, L. (2004). The Amharic dialects revisited. *From beyond the Mediterranean. Akten des*, 7, 449–480.
- Zewde, B. (2002). Systems of local governance among the Gurage. In *The challenge of democracy from below* (pp. 17–28). Nordiska Afrikainstitutet, Uppsala.
- Zobl, H. (1984). The wave model of linguistic change and the naturalness of inter-language. *Studies in Second Language Acquisition*, 6(2), 160–185.

---

# Appendix A

## Data Gathering Tools

### A.1 Background questionnaire

*Purpose: It was employed to determine the students' language background*

Dear students, we use this questionnaire to gather information about the languages which are spoken in your area. Your responses will be kept confidential and are used only for research purposes. Hence, please respond as honest as you can. Thank you for the time you take to fill in the questionnaire.

**Part I: Personal Background**

1. What is your date of birth (DD/MM/YY)? .....
2. Are you a. Male b. Female (circle one)
3. Where is your place of birth (town/village)? .....
4. Where is your present address (town/village)? .....
5. What is your grade level? (Circle correct choice)  
a. Grade nine c. Grade ten d. Grade eleven f. Grade twelve

**Part II: Language Background**

1. What is your first language? \_\_\_\_\_
2. What languages do you speak other than your first language?  
\_\_\_\_\_
3. Which language (s) do your parents speak?  
a) Your Father \_\_\_\_\_  
b) Your mother \_\_\_\_\_
4. Which language is frequently spoken by your friends? \_\_\_\_\_
5. Has your family changed their place of residence? Please indicate the places they lived and the language spoken in each place  

<u>Place</u>	<u>Language</u>
a. _____	_____
b. _____	_____
c. _____	_____
d. _____	_____
6. How often do you use your mother tongue? \_\_\_\_\_  
A. Very often B. Often C. Sometime D. Rarely E. Not at all
7. Which language (s) is spoken in most of the schools you have attended?  
\_\_\_\_\_
8. Which other language is spoken in your vicinity?  
\_\_\_\_\_



## A.2 Response sheet for Words Categorization test

**Instruction:** Dear student, you are going to listen to some list of words. Listen carefully and determine in which of the following categories each word belongs. For one word there is only one possible category. Provide your answer by putting 'X' mark in the box provided in front of each category. Note that for every audio stimulus, there are 10 options of word categories.

1. Cloths
2. Body Parts
3. Kitchen Utilities
4. Fruits
5. Food Type
6. Domestic Animals
7. Furniture
8. Vegetables
9. Wild Animals
10. Cereals

## A.3 Response sheet for perceptual and attitude test

**Direction:** Dear student, you will be presented with ten successive stories. Listen attentively to each of the stories and rate the story based on the questions which are provided below.

1. To what extent do you understand the speaker in the recording? Respond by putting 'X' mark on one of the numbers provided.

Do not understand	0	1	2	3	4	5	6	7	8	9	10	Completely Understand
-------------------	---	---	---	---	---	---	---	---	---	---	----	-----------------------

2. To what extent the recording is similar to your own language? Respond by putting 'X' on one of the numbers.

Not Similar	0	1	2	3	4	5	6	7	8	9	10	Completely Similar
-------------	---	---	---	---	---	---	---	---	---	---	----	--------------------

3. Is the speech of the speaker in the recording beautiful or not compared to your own language? Respond by putting 'X' mark on one of the numbers provided.

Not Beautiful	0	1	2	3	4	5	6	7	8	9	10	Beautiful
---------------	---	---	---	---	---	---	---	---	---	---	----	-----------

# Appendix B

## Test Materials

### B.1 List of words for phonetic and lexical distance

No	English	Amharic	Oromo	Chaha	Endegagn	Ezha	Gumer	Gura	Inor	Kistane	Meskan	Muher	Silt'e
1	added	dämmərə <sup>1</sup>	dabale <sup>1</sup>	dəpərəm <sup>1</sup>	dappərə <sup>1</sup>	dabbərəm <sup>1</sup>	dappərəm <sup>1</sup>	dəpərəm <sup>1</sup>	dappərə <sup>1</sup>	dabbələm <sup>1</sup>	dəbbərəm <sup>1</sup>	dəbbərəm <sup>1</sup>	dabələ <sup>1</sup>
2	all	hullum <sup>1</sup>	hunda <sup>1</sup>	innim <sup>1</sup>	hin?edi <sup>2</sup>	innim <sup>1</sup>	innim <sup>1</sup>	innim <sup>1</sup>	innə <sup>1</sup>	kullim <sup>1</sup>	innim <sup>1</sup>	innim <sup>1</sup>	hullinkə <sup>1</sup>
3	ape	tofa	kamale <sup>2</sup>	wənkə <sup>1</sup>	wenkə <sup>1</sup>	wənkə <sup>1</sup>	wənkə <sup>1</sup>	wankə <sup>1</sup>	wankə <sup>1</sup>	kəmole <sup>2</sup>	kəmole <sup>2</sup>	kəmole <sup>2</sup>	kəmole <sup>2</sup>
4	arm	kind <sup>1</sup>	irre <sup>3</sup>	agat <sup>2</sup>	agad <sup>2</sup>	hinno <sup>1</sup>	xinə <sup>1</sup>	hinə <sup>1</sup>	hinə?ə <sup>1</sup>	kirre <sup>3</sup>	hinno <sup>1</sup>	hitte <sup>1</sup>	kirre <sup>3</sup>
5	back	gərba	duuba	g'isa <sup>1</sup>	g'isa <sup>1</sup>	gini <sup>3</sup>	f'ər <sup>2</sup>	g'isa <sup>1</sup>	gisi?ə <sup>1</sup>	gini <sup>3</sup>	gini <sup>3</sup>	gini <sup>3</sup>	cin <sup>3</sup>
6	barley	gəbs <sup>2</sup>	garbuu <sup>2</sup>	əxir <sup>1</sup>	əhir <sup>1</sup>	əxir <sup>1</sup>	əhir <sup>1</sup>	əhir <sup>1</sup>	əhir <sup>1</sup>	gəbs <sup>2</sup>	əhir <sup>1</sup>	əxi <sup>1</sup>	ixil <sup>1</sup>
7	basil	bəssobila <sup>1</sup>	maslobbaa <sup>1</sup>	məsobila <sup>1</sup>	məsobila <sup>1</sup>	məsobila <sup>1</sup>	bəssobila <sup>1</sup>	bəssobila <sup>1</sup>	məsobila <sup>1</sup>	məsobila <sup>1</sup>	k'asanət <sup>3</sup>	bəssobila <sup>1</sup>	bəkər <sup>2</sup>
8	bean	adang <sup>w</sup> arre <sup>1</sup>	ašongoree <sup>1</sup>	adeng <sup>w</sup> alle <sup>1</sup>	adeng <sup>w</sup> alle <sup>1</sup>	adeng <sup>w</sup> alle <sup>1</sup>	adəng <sup>w</sup> are <sup>1</sup>	adəng <sup>w</sup> aro <sup>1</sup>		bolokke <sup>2</sup>	adəng <sup>w</sup> arre <sup>1</sup>	adəngurro <sup>1</sup>	bolokke <sup>2</sup>
9	beautiful	kongo	bareeddu <sup>2</sup>	mərkama <sup>1</sup>	mərkama <sup>1</sup>	mərkama <sup>1</sup>	mərkama <sup>1</sup>	mərkama <sup>1</sup>	mərkama <sup>1</sup>	mərkama <sup>1</sup>	mərkama <sup>1</sup>	mərkama <sup>1</sup>	bəreedə <sup>2</sup>
10	bed	alga <sup>1</sup>	sire	alga <sup>1</sup>	anz <sup>1</sup>	alga <sup>1</sup>	alga <sup>1</sup>	arga <sup>1</sup>	arga <sup>1</sup>	alga <sup>1</sup>	alga <sup>1</sup>	alga <sup>1</sup>	dugmalə <sup>2</sup>
11	began	gəmmərə <sup>2</sup>	eegale	kənasəm <sup>1</sup>	kənasə <sup>1</sup>	kənasəm <sup>1</sup>	kənasəm <sup>1</sup>	kənasə <sup>1</sup>	kənasə <sup>1</sup>	kərrəso <sup>1</sup>	kərrəso <sup>1</sup>	kərrəso <sup>1</sup>	gəmmərə <sup>2</sup>
13	belt	kəbatto <sup>1</sup>	kəbattoo <sup>1</sup>	mətagəca <sup>2</sup>	kəbatto <sup>1</sup>	kəbatto <sup>1</sup>	kəbatto <sup>1</sup>	kəbatto <sup>1</sup>	zələbat <sup>3</sup>	kəbatto <sup>1</sup>	kəbatto <sup>1</sup>	kəbatto <sup>1</sup>	kəbatto <sup>1</sup>
14	basso	bəso <sup>1</sup>	basso <sup>1</sup>	bəso <sup>1</sup>	bəso <sup>1</sup>	bəso <sup>1</sup>	bəso <sup>1</sup>	bəso <sup>1</sup>	bəso <sup>1</sup>	bəso <sup>1</sup>	bəso <sup>1</sup>	bəso <sup>1</sup>	bəso <sup>1</sup>
15	big	tillik	guddaa	nikiya <sup>1</sup>	nu <sup>2</sup>	imiyə <sup>1</sup>	nikiya <sup>1</sup>	nikiya <sup>1</sup>	nu?iya <sup>1</sup>	malak <sup>2</sup>	nik <sup>1</sup>	li?iya <sup>1</sup>	yəroore <sup>3</sup>
16	black	t'ukur <sup>1</sup>	gurraacča	t'ukur <sup>1</sup>	gombəna <sup>2</sup>	gombəna <sup>2</sup>	t'ukur <sup>1</sup>	t'ukur <sup>1</sup>	gombəna <sup>2</sup>	t'ukur <sup>1</sup>	t'ukur <sup>1</sup>	t'ukur <sup>1</sup>	t'em <sup>3</sup>
17	blew	nəffəsa <sup>1</sup>	afuufe	nəffəsəm <sup>1</sup>	nəffəsə <sup>1</sup>	nəffəsəm <sup>1</sup>	nəffəsəm <sup>1</sup>	nəffəsəm <sup>1</sup>	nəffəsə <sup>1</sup>	nəffəsə <sup>1</sup>	nəffəsə <sup>1</sup>	nəffəsəm <sup>1</sup>	nəffəsə <sup>1</sup>
18	blood	dəm <sup>1</sup>	diga	dəm <sup>1</sup>	dəm <sup>1</sup>	dəm <sup>1</sup>	dəm <sup>1</sup>	dəm <sup>1</sup>	dəm <sup>1</sup>	dəm <sup>1</sup>	dəm <sup>1</sup>	dəm <sup>1</sup>	dəm <sup>1</sup>
19	boiled	afolla <sup>2</sup>	danfise	čəkərəm <sup>1</sup>	čakarra <sup>1</sup>	čəkərə <sup>1</sup>	čəkərəm <sup>1</sup>	čəkərə <sup>1</sup>	čəkərə <sup>1</sup>	fəllam <sup>2</sup>	čəkərə <sup>1</sup>	čəkərəm <sup>1</sup>	fəla <sup>2</sup>
20	bone	afim <sup>1</sup>	lafee	afim <sup>1</sup>	afim <sup>1</sup>	afim <sup>1</sup>	afim <sup>1</sup>	afim <sup>1</sup>	afim <sup>1</sup>	afim <sup>1</sup>	afim <sup>1</sup>	afim <sup>1</sup>	hafim <sup>1</sup>
21	bread	dabbo <sup>1</sup>	daabbo <sup>1</sup>	dabbo <sup>1</sup>	dabo?ə <sup>1</sup>	dabbo <sup>1</sup>	dabbo <sup>1</sup>	dabbo <sup>1</sup>	dabbo?ə <sup>1</sup>	dabbo <sup>1</sup>	furno <sup>2</sup>	dabbo <sup>1</sup>	furno <sup>2</sup>
22	breakfast	kurs	čiree	ginzir <sup>1</sup>	ginzir <sup>1</sup>	ginzir <sup>1</sup>	ginzir <sup>1</sup>	ginzir <sup>1</sup>	ginzir <sup>1</sup>	ginzir <sup>1</sup>	yaddara <sup>2</sup>	addara <sup>2</sup>	ginzir <sup>1</sup>
23	breast	t'ut <sup>1</sup>	harma	t'ut <sup>1</sup>	t'uw <sup>1</sup>	t'uw <sup>1</sup>	t'ut <sup>1</sup>	t'ut <sup>1</sup>	t'ut <sup>1</sup>	t'ibuyya <sup>1</sup>	t'ub <sup>1</sup>	t'ibuyya <sup>1</sup>	t'ub <sup>1</sup>
24	brother	wəndimm	obboleessa	g'əpa <sup>1</sup>	əšša?əm <sup>2</sup>	g'əbbe <sup>1</sup>	g'əppay <sup>1</sup>	g'əpa <sup>1</sup>	əsam <sup>2</sup>	zamm <sup>2</sup>	g'əbbe <sup>1</sup>	g'əbbe <sup>1</sup>	indot <sup>1</sup>
25	bull	bərə <sup>2</sup>	dibičča	wir <sup>1</sup>	wir <sup>1</sup>	wir <sup>1</sup>	wir <sup>1</sup>	wir <sup>1</sup>	wir <sup>1</sup>	wir <sup>1</sup>	wir <sup>1</sup>	borə <sup>2</sup>	wir <sup>1</sup>
26	cabbage	gommənn	raafu	ambir <sup>1</sup>	ambir <sup>1</sup>	ambir <sup>1</sup>	ambir <sup>1</sup>	ambir <sup>1</sup>	ambir <sup>1</sup>	ambir <sup>1</sup>	ambir <sup>1</sup>	ambir <sup>1</sup>	hamil <sup>1</sup>
27	calf (n)		g'abbi	m'əsa <sup>1</sup>	mosa <sup>1</sup>	m'əssa <sup>1</sup>	m'əsa <sup>1</sup>	m'əsa <sup>1</sup>	m'əsa <sup>1</sup>	t'əg <sup>2</sup>	m'asa <sup>1</sup>	dəg <sup>2</sup>	izək <sup>3</sup>
28	came	məf'aa <sup>2</sup>	dufe	čəna <sup>1</sup>	mə?a <sup>2</sup>	čəna <sup>1</sup>	čəna <sup>1</sup>	čəna <sup>1</sup>	čəna <sup>1</sup>	mə?a <sup>2</sup>	məf'aa <sup>2</sup>	čəna <sup>1</sup>	bəssam <sup>3</sup>
29	came out	wəf'aačč <sup>1</sup>	baate	wəf'am <sup>1</sup>	wə?ačč <sup>1</sup>	wəf'aaččim <sup>1</sup>	wəf'aaččim <sup>1</sup>	wəf'aaččim <sup>1</sup>	wə?ačč <sup>1</sup>	wəf'atti <sup>1</sup>	wəf'aa <sup>2</sup>	wəf'aaččim <sup>1</sup>	wəf'at <sup>1</sup>
30	camel	giməl <sup>1</sup>	gaala <sup>1</sup>	gamera <sup>1</sup>	giməl <sup>1</sup>	gamera <sup>1</sup>	gamera <sup>1</sup>	gamera <sup>1</sup>	gamera <sup>1</sup>	giməl <sup>1</sup>	gamelə <sup>1</sup>	gamelə <sup>1</sup>	gaameela <sup>1</sup>
31	cat	dimmət	adurre <sup>2</sup>	angača <sup>1</sup>	angaččə <sup>1</sup>	angačča <sup>1</sup>	angačča <sup>1</sup>	angača <sup>1</sup>	angača <sup>1</sup>	angača <sup>1</sup>	angača <sup>1</sup>	angača <sup>1</sup>	aadən <sup>2</sup>
32	chair	wənbər <sup>1</sup>	barčuma <sup>4</sup>	wənbər <sup>1</sup>	t'əkəša <sup>2</sup>	wənbər <sup>1</sup>	wənbər <sup>1</sup>	wənbər <sup>1</sup>	t'əkəša <sup>2</sup>	wənbər <sup>1</sup>	sorər <sup>3</sup>	wənbər <sup>1</sup>	borčuma <sup>4</sup>
33	chassis	rəkəbot <sup>1</sup>	rakaboot <sup>2</sup>	yəsinəčə <sup>1</sup>	rəkəbot <sup>1</sup>	yəsinəččə <sup>1</sup>	yəsinəččə <sup>1</sup>	yəsinəččə <sup>1</sup>	sine? <sup>1</sup>	rəkəbot <sup>1</sup>	yəsinəččə <sup>1</sup>	yəsinəččə <sup>1</sup>	gənəč <sup>3</sup>
34	cheap	rikkaš <sup>3</sup>	rakiša <sup>3</sup>	giwa <sup>1</sup>	wuda? <sup>2</sup>	giwa <sup>1</sup>	giwa <sup>1</sup>	giwa <sup>1</sup>	widi? <sup>2</sup>	arkus <sup>3</sup>	rikkaš <sup>3</sup>	rikkaš <sup>3</sup>	riks <sup>3</sup>
35	cheek	gunc <sup>2</sup>	mallaa	danga <sup>1</sup>	danga <sup>1</sup>	danga <sup>1</sup>	danga <sup>1</sup>	danga <sup>1</sup>	guncə <sup>2</sup>	guncə <sup>2</sup>	danga <sup>1</sup>	guncə <sup>2</sup>	guncə <sup>2</sup>
36	chest	dərət <sup>1</sup>	lap'pe	data <sup>1</sup>	datta <sup>1</sup>	dadda <sup>1</sup>	datta <sup>1</sup>	fanka <sup>2</sup>	data <sup>1</sup>	dərət <sup>1</sup>	lin <sup>3</sup>	dadda <sup>1</sup>	wəzənə <sup>3</sup>
37	chickpea	šimbura <sup>1</sup>	šumburaa <sup>1</sup>	čəmb'əra <sup>1</sup>	šumbura <sup>1</sup>	čəmb'əra <sup>1</sup>	šimbura <sup>1</sup>	čimbura <sup>1</sup>	šumbura <sup>1</sup>	šimbura <sup>1</sup>	šimbura <sup>1</sup>	šimbura <sup>1</sup>	šimbura <sup>1</sup>
38	child	hišan	daa?ma	tika <sup>1</sup>	tika <sup>1</sup>	tika <sup>1</sup>	tika <sup>1</sup>	tika <sup>1</sup>	tika <sup>1</sup>	bayy <sup>2</sup>	tika <sup>1</sup>	tika <sup>1</sup>	čilo <sup>3</sup>
39	cloak(n)	kot <sup>1</sup>	koota <sup>1</sup>	kot <sup>1</sup>	kot <sup>1</sup>	kot <sup>1</sup>	kot <sup>1</sup>	kot <sup>1</sup>	kot <sup>1</sup>	kot <sup>1</sup>	kot <sup>1</sup>	kot <sup>1</sup>	koote <sup>1</sup>
40	cloud	damməna <sup>1</sup>	duumessa <sup>1</sup>	dabəra <sup>1</sup>	dawəna <sup>1</sup>	dabəra <sup>1</sup>	dabəra <sup>1</sup>	dabəra <sup>1</sup>	daməra <sup>1</sup>	daməna <sup>1</sup>	dabəna <sup>1</sup>	dabəna <sup>1</sup>	dabəna <sup>1</sup>







111	insult (v)	səddəbə²	arrabse	kəndməᵐ¹	kənəmə¹	kənəməmᵐ¹	kīnema¹	kəndməᵐ¹	kənəmə¹	səddəbəm²	səddəbə²	səddəbəm¹	sədəbə²
112	job/work	sira	hogū	mena¹	mena¹	mena¹	mena¹	mena¹	mena¹	wəzələ²	menna¹	merra¹	bū³
113	joke (n)	kəld¹	keesa¹	kəṣə¹	kəld¹	kəṣə¹	k'əṣə¹	kəld¹	siya²	kəld¹	kəld¹	kəld¹	kəld¹
114	kettle	gəbənā¹	gəbana¹	gəbən¹	gəbən¹	gəbən¹	gəbən¹	gəbən¹	gəbəlʔən¹	gəbana¹	gəbən¹	gəbana¹	gəbən¹
115	king	nigus¹	nigusa¹	nig'is¹	nigus¹	nig'is¹	nigus¹	nig'is¹	nig'is¹	nigus¹	nig'is¹	nigus¹	nigus¹
116	knife	billəwa	able	sənda³	ṣəṭta²	sənda³	sənda³	wəkara¹	ṣəṭa²	golod⁴	golodo⁴	sənda³	golodo⁴
117	Koććo	k'ocćo	k'ooććoo	wissa¹	kusa¹	wissa¹	wissa¹	wisa¹	wisʔa¹	əkusa¹	wissa¹	biya²	finaniććo³
118	ladle	čilfa¹	čilfaa¹	čilfa¹	čilfa¹	čilfa¹	čilfa¹	nikankəfwə²	ankəfo²	čilfa¹	čilfa¹	čilfa¹	čilfa¹
119	leg	igir¹	milla	əgir¹	əgir¹	igir¹	əgir¹	əgir¹	əgir¹	əgir¹	əgir¹	əgir¹	ingir¹
120	lemon	lomi¹	loomi¹	lomi¹	lomi¹	lomi¹	lomi¹	lomi¹	lomin¹	lomi¹	lomi¹	lomi¹	lomi¹
121	lentil	missir²	missira²	kəsəm¹	kəsəm¹	kəsəm¹	kəsəm¹	kəsəm¹	kəsəm¹	mišira²	kəsəm¹	kəsəm¹	mišir²
122	lion	ambəssa²	leencə	ṣəp¹	ambəssa²	ambəssa²	ṣəpp¹	ṣəpə¹	ṣəp¹	ambəssa²	ambəssa²	ambəssa²	wərba³
123	lip	kənfər¹	hidū	kənfər¹	kənfər¹	kənfər¹	kənfər¹	amz²	kənfər¹	kənfər¹	kənfər¹	kənfər¹	gir'ce³
124	louse	kimal¹	ingire	kimar¹	iwan¹	kimar¹	kimar¹	kimar¹	imar¹	kimal¹	kimal¹	kime¹	kumal¹
125	love	fikur²	gəalala	nimağə¹	nimağə¹	nimağə¹	nimağə¹	nimağə¹	aʔimaga¹	fikir²	dad²	widdan⁴	dəğ³
126	maize	bək'ollo¹	bək'oollo¹	bək'ollo¹	bək'okolo¹	bək'ikwollo¹	bək'wollo¹	bək'ollo¹	bək'ollo¹	bək'ollo¹	bək'ollo¹	bək'ollo¹	bək'ollo¹
127	man	səw¹	nama	səb¹	səw¹	səb¹	mis²	mis²	səb¹	mis²	mis²	mis²	səb¹
128	mattress	firas²	firaasa²	kap'at¹	kappad¹	kabb'at¹	firas²	kabbat¹	kap'at¹	firas²	firas²	kabb'at¹	kabbat¹
129	means	biləhat¹	mala²	ayo²	bilat¹	bilat¹	bilat¹	ayu²	beyə¹	bilat¹	bilat¹	biləhat¹	mala²
130	mirror	məstawat¹	daawiti	məsəwət¹	məstawəd¹	məsərewat¹	məstawat¹	məstawət¹	məstawəd¹	məstawat¹	məstawat¹	məstawat¹	məstawat¹
131	monday	səñho	wit'ət¹	wit'ət¹	wit'ət¹	wit'ət¹	wit'ət¹	wit'ət¹	wit'ət¹	wit'ət¹	wit'ət¹	wit'ət¹	wit'ət¹
132	monkey	zing'oro¹	gəaldeessa	zangera¹	zangəro¹	zang'oro¹	zang'oro¹	zang'oro¹	zang'oro¹	zang'oro¹	zang'oro¹	zang'oro¹	zang'oro¹
133	more	bəʔam³	daral	nūkar¹	muʔar¹	nūkar¹	nūkar¹	nūkar¹	buseʔəho²	burim³	bəʔam³	bəʔam³	yəbəza⁴
134	mule	bək'lo¹	gaange	bik'ərə¹	bakurə¹	buk'ərə¹	bikurə¹	bikurə¹	bukurə¹	bicil¹	b'ək'ila¹	buʔurə¹	bok'lo¹
135	mush	nifro¹	šummo	nifro¹	buṣəʔa²	čəkoreṭ²	nif'ro¹	nifro¹	wifro¹	nifon¹	čik'ərə³	čəkoran³	inšikora³
136	nail	tifir¹	keensa	tifir¹	infir¹	tifir¹	tifir¹	tifir¹	infir¹	təfir¹	tifir¹	tifir¹	tifir¹
137	neck	angət¹	morma	angət¹	angət¹	angət¹	angət¹	angət¹	angəd¹	angət¹	angət¹	angət¹	ang'oro¹
138	niger seeds	nug¹	nuugi¹	nug¹	nug¹	nug¹	nug¹	nug¹	nug¹	nug¹	nug¹	nug¹	nug¹
139	nine	zəʔəñ¹	sagal	ṣəʔə¹	ṣiʔa¹	ṣəʔə¹	ṣəʔə¹	ṣəʔə¹	ṣəʔə¹	ṣəʔəñ¹	ṣəʔə¹	ṣəʔə¹	zəʔ'əñe¹
140	now	ahum¹	amma	əx' a¹	akkə¹	əx' a¹	əx' a¹	əx' a¹	waka¹	axu¹	əx' a¹	axuñña¹	akku¹
141	oat	agğə¹	agğəa¹	agğə¹	humboliʔə²	imboriyat²	agğə¹	agğə¹	agğə¹	imbor²	agğə¹	agğə¹	agğə¹
142	one	and¹	tokko	at¹	att¹	att¹	at¹	at¹	at¹	att¹	att¹	att¹	aad¹
143	onion	šinkurt¹	šunkurti¹	šink'ort¹	šinkurta¹	šink'ort¹	šinkurt¹	šinkurt¹	šink'ort¹	šinkurt¹	šinkurt¹	šinkurt¹	šinkurt¹
144	orange	birtuk' an¹	burtukaana¹	birtuk' an¹	b'irtukan¹	birtuk' an¹	birtukan¹	birtuk' an¹	birtuk' an¹	birtukan¹	birtukan¹	birtukan¹	birtukan¹
145	our	yəñña¹	keēñña¹	yina¹	inay¹	yina¹	yina¹	yina¹	yəñña¹	yəñña¹	yəñña¹	yəñña¹	yəñña¹
146	ox	bərə¹	dibiććə	bora¹	bawra¹	bora¹	bora¹	bora¹	bawərə¹	bora¹	bora¹	bora¹	kəraab²
147	pan	dist¹	dist¹	dist¹	dist¹	dist¹	dist¹	dist¹	dist¹	dist¹	dist¹	dist¹	dist¹
148	pancake	kitt'a³	kitt'aa³	t'irəwiṣə¹	t'orəṣə¹	məʔmo²	kitt'aa³	t'irəwiṣə¹	t'orəwiṣə¹	gums⁴	kitt'a³	kitt'a³	t'oroṣ⁵
149	pea	atar¹	atar¹	getərə¹	gite¹	getərə¹	getərə¹	getərə¹	getərə¹	kiššuwə²	atar¹	getərə¹	gite¹
150	peach	kok¹	kooki¹	kok¹	kok¹	kok¹	kok¹	kok¹	kok¹	kok¹	kok¹	kok¹	kok¹
151	pepper	kariya¹	kür'caa	kare¹	kareṭ²	kale¹	kariya¹	kari¹	kondeʔə²	kariya¹	kariya¹	kare¹	turga²
152	pillow	tiras¹	boraati	tiras¹	tiras¹	tiras¹	tiras¹	gimma²	gimʔə²	gimma²	təras¹	tiras¹	gimma²
153	plate	saxin¹	sahaana¹	saxin¹	sahin¹	disko²	sahin¹	sahin¹	sahan¹	sahan¹	təba²	sahin¹	saan¹
154	porridge	gəñfo³	markaa	ozat¹	daʔə²	ozat¹	ozat¹	owzat¹	daʔə²	gəñfo³	in'kaććə⁴	ozat¹	in'kaććə⁴
155	potato	dinnic'č¹	dinnic'čaa¹	dini'ča¹	dini'čə¹	dinnic'čaa¹	dinnic'čaa¹	dinnic'čaa¹	dinnic'čaa¹	dinnic'čaa¹	dinnic'čaa¹	dinnic'čaa¹	dinnic'čaa¹
156	pumpkin	dubba¹	dubbaa¹	dəbak'ila¹	dəbak'ula¹	dəbak'ulla¹	dəbak'ulla¹	dəwak'ula¹	t'əbak'ula¹	dəbak'ulla¹	dəbak'ulla¹	dəwak'ulla¹	dəbak'ulla¹
157	rain	zinaᵇ¹	rooba	zirab¹	diya²	zirab¹	diya²	zinab¹	diya²	zinab¹	zinab¹	zinab¹	zilaᵇ¹
158	red	kəyy	dūmaa	biṣə¹	buṣə¹	biṣə¹	biṣsa¹	biṣə¹	biṣə¹	biṣsa¹	biṣsa¹	biṣsa¹	buṣə¹
159	respected	akabbərə	kabag'e¹	kəbərəᵐ¹	həbbada²	kəbbərəᵐ¹	təhəbədəᵐ¹	kəbərə¹	xəbədə²	akabbərəᵐ¹	kəbbərə¹	kəbbərəᵐ¹	kəbərə¹
160	rice	ruz¹	ruza¹	ruz¹	ruz¹	ruz¹	ruz¹	ruz¹	ruz¹	ruz¹	ruz¹	ruz¹	ruz¹
161	rich	habtam	dureessa²	dəngəñña¹	dəngəñña¹	dəngəñña¹	dəngəñña¹	dəngəñña¹	dəngəñña¹	dəngəñña¹	dəngəñña¹	dəngəñña¹	duressa²
162	roasted grain	kolo¹	hunikaa	k'wərə¹	kərə¹	k'wərə¹	k'wərə¹	k'wərə¹	k'wərə¹	k'wərə¹	k'wərə¹	k'wərə¹	inkolo¹
163	road	məngəd	amma¹	ema¹	meya¹	ema¹	ema¹	ema¹	meʔa¹	moćə²	ema¹	zəba³	ungə²
164	round	zuriya¹	naannawa	innim²	bihahiʔedi³	zurata¹	zuriya¹	xipibar⁴	zuriya¹	zuriya¹	anke⁵	zuriya¹	zuura¹
165	said	alo¹	gəde	barəᵐ¹	bara¹	barəᵐ¹	barəᵐ¹	barə¹	barə¹	balo¹	barə¹	bem¹	baala¹
166	salt	čəw	soogida³	asso¹	asso¹	asso¹	asso¹	asso¹	bəʔəd²	soğida³	asso¹	assəwa¹	arus¹
167	saturday	kidame	sanbata¹	kəʔansənbat¹	ennəgəya²	kəʔansənbat¹	kəʔansənbat¹	kəʔansənbat¹	kəʔansənbat¹	kidansənbat¹	kidansənbat¹	kəʔansənbat¹	ansənbat¹
168	saw	ayya¹	arge	aṣə¹	aššəm¹	aṣə¹	aššəm¹	aṣə¹	aṣə¹	aṣə¹	aṣə¹	aṣə¹	anəṣ¹
169	seed	zər¹	šəñni	zər¹	zəʔə¹	zər¹	zər¹	zər¹	zər¹	zər¹	zər¹	zər¹	zər¹
170	seven	səbat¹	torba	səbat¹	səwʔat¹	səbat¹	səbat¹	səbat¹	saat¹	səbat¹	sebat¹	səbat¹	saʔabat¹
171	she	is' a³	isə³	x'it¹	šido²	x'it¹	x'it¹	x'it¹	xida¹	kiya¹	xiti¹	x' a³	iša³
172	sheep	bəg	hoolaa	t'ə¹	t'ay¹	t'ay¹	t'ay¹	t'ə¹	t'ay¹	t'ay¹	t'ə¹	t'ə¹	t'ay¹
173	shoe	čamma¹	kop'ee²	čamma¹	čamma¹	čamma¹	čamma¹	čamma¹	čamma¹	kobe²	čamma¹	čamma¹	čamma¹
174	short	ac'cir¹	gabaabaa	ac'ir¹	iʔir¹	ac'ir¹	ac'ir¹	ac'ir¹	eʔir¹	ac'cir¹	ac'cir¹	ac'cir¹	anc'ir¹
175	short trouser	kumf'a¹	kumf'aa¹	kumf'a¹	kumf'a¹	kumf'a¹	kumf'a¹	kumf'a¹	končaru²	kumf'a¹	bogge³	kumf'a¹	bogge³
176	six	siddist¹	gəha	siddist¹	siddist¹	siddist¹	siddist¹	siddist¹	siddist¹	siddist¹	siddist¹	siddist¹	siddist¹
177	slowly	bəkəssita²	suuttatti	təhin¹	kəsbərta²	təhin¹	təx'ix'm¹	təx'ix'm¹	təhin¹	digbəwəbal²	digbərəᵐ³	diggebem³	balloḍit⁴
178	smoke	čis	aara	tən¹	tan¹	tən¹	tən¹	tən¹	tən¹	tən¹	tən¹	tən¹	tən¹
179	snatched	nəʔ'əkə³	bute	məććəkəᵐ¹	nəʔ'əkə¹	nəʔ'əkə¹	nəʔ'əkəᵐ¹	nəʔ'əkə¹	nəʔ'əkəᵐ¹	məććəkəᵐ¹	məććəkəᵐ¹	nəʔ'əkəᵐ¹	boc'ćəkə¹
180	so	siləzzih⁵	kanaafu	ikka¹	hənəṭa²	yəxər³	yəxare³	əxir³	buʔuktu⁴	siləzza⁵	siləhana⁵	siləzzih⁵	loonəmko⁶

181	song	zəfən	sirba	lale <sup>1</sup>	das <sup>2</sup>	lale <sup>1</sup>	weg <sup>3</sup>	lale <sup>1</sup>	lale <sup>1</sup>	dərs <sup>2</sup>	dərs <sup>2</sup>	dərris <sup>4</sup>	šibat <sup>5</sup>
182	soon	wadiyaw <sup>4</sup>	battalumatti	himtəhm <sup>1</sup>	hamgizzi <sup>2</sup>	hingiz <sup>2e</sup>	ximtaxim <sup>1</sup>	kiməga <sup>2</sup>	himəga <sup>2</sup>	zagizze <sup>2</sup>	kəmo <sup>3</sup>	wədiyaw <sup>4</sup>	waktəy <sup>5</sup>
183	sorghum	masilla <sup>1</sup>	bisingaa	masilla <sup>1</sup>	masulla <sup>1</sup>	masilla <sup>1</sup>	masilla <sup>1</sup>	masilla <sup>1</sup>	m <sup>1</sup> asina <sup>1</sup>	masilla <sup>1</sup>	masilla <sup>1</sup>	masilla <sup>1</sup>	masulle <sup>1</sup>
184	steer	wəfən <sup>1</sup>	waatiyyo	wəfer <sup>1</sup>	wənfir <sup>1</sup>	wəfir <sup>1</sup>	wəfer <sup>1</sup>	wəfir <sup>1</sup>	wənfir <sup>1</sup>	mofen <sup>1</sup>	wəfm <sup>1</sup>	wəfen <sup>1</sup>	mofən <sup>1</sup>
185	stew	wət <sup>1</sup>	ittoo	wət <sup>1</sup>	wət <sup>1</sup>	wət <sup>1</sup>	wət <sup>1</sup>	wət <sup>1</sup>	wət <sup>1</sup>	wət <sup>1</sup>	wət <sup>1</sup>	wət <sup>1</sup>	wət <sup>1</sup>
186	stirring rod	mamasəya <sup>2</sup>	kokorsaa	wəgəngə <sup>1</sup>	mammase <sup>2</sup>	wəgəngə <sup>1</sup>	wəgəngə <sup>1</sup>	wəgəngə <sup>1</sup>	wəgəngə <sup>1</sup>	wəgəngə <sup>1</sup>	wəlgəgə <sup>1</sup>	wəlgəgə <sup>1</sup>	hunkutma <sup>3</sup>
187	stone	dingay	dagaa	imir <sup>1</sup>	əwn <sup>2</sup>	imir <sup>1</sup>	imir <sup>1</sup>	imir <sup>1</sup>	imir <sup>1</sup>	yimir <sup>1</sup>	imir <sup>1</sup>	imafina <sup>1</sup>	uum <sup>2</sup>
188	stool	bərcuma <sup>1</sup>	bərcuma <sup>1</sup>	b <sup>w</sup> ərcuma <sup>1</sup>	bərcumma <sup>1</sup>	b <sup>w</sup> ərcuma <sup>1</sup>	borcuma <sup>1</sup>	wərcuma <sup>1</sup>	bərcuma <sup>1</sup>	wənbər <sup>2</sup>	borcuma <sup>1</sup>	b <sup>w</sup> ərcimma <sup>1</sup>	kattə <sup>3</sup>
189	stronger	fənkarra <sup>4</sup>	irra cimaa	gədana <sup>1</sup>	tuʔə <sup>2</sup>	tiwwə <sup>2</sup>	tuwə <sup>2</sup>	tiwwə <sup>2</sup>	tuʔə <sup>2</sup>	bərtu <sup>1</sup>	bərtu <sup>3</sup>	fənkarra <sup>4</sup>	gədit <sup>1</sup>
190	sucked	fəbba <sup>1</sup>	luuge	təp <sup>w</sup> əm <sup>1</sup>	fəppə <sup>1</sup>	fəbb <sup>w</sup> əm <sup>1</sup>	fəpp <sup>w</sup> əm <sup>1</sup>	təp <sup>w</sup> əm <sup>1</sup>	fəp <sup>w</sup> ə <sup>1</sup>	fəbbəm <sup>1</sup>	fəbba <sup>1</sup>	fəbb <sup>w</sup> əm <sup>1</sup>	tobe <sup>1</sup>
191	suddenly	dingət <sup>2</sup>	tasa	matəga <sup>1</sup>	immatəg <sup>1</sup>	immatigi <sup>1</sup>	dingət <sup>2</sup>	himəga <sup>1</sup>	dingət <sup>2</sup>	dingət <sup>2</sup>	dingət <sup>2</sup>	dingət <sup>2</sup>	dinbut <sup>2</sup>
192	sun	šəhay	aduu <sup>2</sup>	čet <sup>1</sup>	ayed <sup>2</sup>	čet <sup>1</sup>	čet <sup>1</sup>	čet <sup>1</sup>	iwəya <sup>1</sup>	yimir <sup>2</sup>	čet <sup>1</sup>	aret <sup>3</sup>	ayr <sup>2</sup>
193	sunday	ihud	dilbata	wursənbət <sup>1</sup>	abba <sup>2</sup>	wirsənbət <sup>1</sup>	wirsənbət <sup>1</sup>	wursənbət <sup>1</sup>	wursənbət <sup>1</sup>	wursənbət <sup>1</sup>	wursənbət <sup>1</sup>	wursənbət <sup>1</sup>	gidirsənbət <sup>1</sup>
194	sunflower	suf <sup>1</sup>	sufu <sup>1</sup>	suf <sup>1</sup>	suf <sup>1</sup>	suf <sup>1</sup>	suf <sup>1</sup>	suf <sup>1</sup>	suf <sup>1</sup>	suf <sup>1</sup>	suf <sup>1</sup>	suf <sup>1</sup>	suf <sup>1</sup>
195	sweat	lab	hurkaa	wizat <sup>1</sup>	azat <sup>1</sup>	wizat <sup>1</sup>	wizat <sup>1</sup>	lab <sup>2</sup>	azat <sup>1</sup>	wizat <sup>1</sup>	wizat <sup>1</sup>	wizat <sup>1</sup>	wizat <sup>1</sup>
196	takeoff	awəlləkə <sup>2</sup>	baase	watməni <sup>1</sup>	awənəga <sup>2</sup>	af <sup>w</sup> am <sup>1</sup>	aw <sup>w</sup> ot <sup>1</sup>	waw <sup>w</sup> ta <sup>1</sup>	atwanukt <sup>1</sup>	awəlləkə <sup>3</sup>	wawlək <sup>2</sup>	awəlləʔəm <sup>2</sup>	šeelətə <sup>4</sup>
197	tangerine	məndərin <sup>1</sup>	mandarīna <sup>1</sup>	məndəroni <sup>1</sup>	məndərin <sup>1</sup>	məndəroni <sup>1</sup>	məndərin <sup>1</sup>	məndəroni <sup>1</sup>	məndəroni <sup>1</sup>	məndəroni <sup>1</sup>	məndərin <sup>1</sup>	məndəroni <sup>1</sup>	məndəroni <sup>1</sup>
198	teff	fəf <sup>1</sup>	fəafi <sup>1</sup>	fəfi <sup>1</sup>	fəf <sup>1</sup>	fəfi <sup>1</sup>	fəfi <sup>1</sup>	fəfi <sup>1</sup>	fəfi <sup>1</sup>	fəfi <sup>1</sup>	fəfi <sup>1</sup>	fəfi <sup>1</sup>	fəfə <sup>1</sup>
199	ten	asir <sup>1</sup>	kudan	asir <sup>1</sup>	asir <sup>1</sup>	asi <sup>1</sup>	asir <sup>1</sup>	asir <sup>1</sup>	asir <sup>1</sup>	asir <sup>1</sup>	asir <sup>1</sup>	asir <sup>1</sup>	asir <sup>1</sup>
200	that	ya <sup>2</sup>	san	xix <sup>1</sup>	xada <sup>1</sup>	za <sup>2</sup>	xix <sup>1</sup>	hata <sup>1</sup>	xada <sup>1</sup>	za <sup>2</sup>	za <sup>2</sup>	za <sup>2</sup>	hattay <sup>1</sup>
201	then	kazziya	sanbooda	təhankə <sup>1</sup>	təhəʔegad <sup>1</sup>	təhankye <sup>1</sup>	taxankə <sup>1</sup>	təhʔang <sup>w</sup> ə <sup>1</sup>	təhim <sup>1</sup>	tək <sup>w</sup> ə <sup>2</sup>	bahu <sup>3</sup>	təhimgə <sup>4</sup>	təʔizof <sup>5</sup>
202	they	iməssu	isaan	xino <sup>1</sup>	huno <sup>1</sup>	xino <sup>1</sup>	xino <sup>1</sup>	xinu <sup>1</sup>	xinowa <sup>1</sup>	kinnəm <sup>1</sup>	xirno <sup>1</sup>	xinnəw <sup>1</sup>	uxun <sup>1</sup>
203	thick	wəfram	furdaa	gəngir <sup>1</sup>	gušir <sup>1</sup>	gəngir <sup>1</sup>	gəngir <sup>1</sup>	gəngir <sup>1</sup>	gəngir <sup>1</sup>	dadan <sup>2</sup>	gəngin <sup>1</sup>	gəngim <sup>1</sup>	mogarr <sup>3</sup>
204	thief	leba <sup>1</sup>	hattuu	neba <sup>1</sup>	fangəya <sup>2</sup>	səwəni <sup>3</sup>	nəba <sup>1</sup>	neba <sup>1</sup>	fəngəya <sup>2</sup>	leba <sup>1</sup>	leba <sup>1</sup>	leba <sup>1</sup>	rangit <sup>4</sup>
205	thin	kəccin <sup>2</sup>	kallaa	sisə <sup>1</sup>	kəccin <sup>2</sup>	sisso <sup>1</sup>	sisə <sup>1</sup>	sisə <sup>1</sup>	kəccir <sup>2</sup>	as <sup>t</sup> atib <sup>3</sup>	kəccin <sup>2</sup>	səssə <sup>3</sup>	kəccin <sup>2</sup>
206	this	yih <sup>1</sup>	kun	zi <sup>1</sup>	wa <sup>2</sup>	zix <sup>1</sup>	zi <sup>1</sup>	zi <sup>1</sup>	wəda <sup>2</sup>	zi <sup>1</sup>	zi <sup>1</sup>	zi <sup>1</sup>	imnə <sup>3</sup>
207	three	sost <sup>1</sup>	sadi <sup>1</sup>	sost <sup>1</sup>	sost <sup>1</sup>	sost <sup>1</sup>	sost <sup>1</sup>	sost <sup>1</sup>	soʔost <sup>1</sup>	sost <sup>1</sup>	sost <sup>1</sup>	sost <sup>1</sup>	šest <sup>1</sup>
208	thursday	hamus <sup>1</sup>	kamisa <sup>1</sup>	amus <sup>1</sup>	amus <sup>1</sup>	amus <sup>1</sup>	am <sup>w</sup> is <sup>1</sup>	amus <sup>1</sup>	amus <sup>1</sup>	xamus <sup>1</sup>	amus <sup>1</sup>	amus <sup>1</sup>	xamus <sup>1</sup>
209	tiger	nəbir <sup>3</sup>	kərreensa	ʒəg <sup>w</sup> ara <sup>1</sup>	zagora <sup>1</sup>	zəg <sup>w</sup> ara <sup>1</sup>	zəg <sup>w</sup> ara <sup>1</sup>	zəg <sup>w</sup> ara <sup>1</sup>	zagora <sup>1</sup>	buswa <sup>2</sup>	zogara <sup>1</sup>	zogara <sup>1</sup>	nawr <sup>3</sup>
210	till	dirəs <sup>1</sup>	haanga	hihedar <sup>1</sup>	haʔed <sup>2</sup>	dar <sup>1</sup>	dar <sup>1</sup>	dar <sup>1</sup>	sənəga <sup>3</sup>	dirəs <sup>1</sup>	sim <sup>4</sup>	dirəs <sup>1</sup>	gəango <sup>5</sup>
211	time	gizze <sup>1</sup>	yaroo	əga <sup>2</sup>	gidad <sup>2</sup>	əga <sup>2</sup>	gam <sup>w</sup> ə <sup>1</sup>	gize <sup>1</sup>	giziyə <sup>1</sup>	giz <sup>w</sup> e <sup>1</sup>	giz <sup>w</sup> e <sup>1</sup>	ge <sup>1</sup>	wəkt <sup>4</sup>
212	to	lə <sup>3</sup>	f	yə <sup>1</sup>	ʔə <sup>2</sup>	lə <sup>3</sup>	yə <sup>1</sup>	yə <sup>1</sup>	ʔə <sup>2</sup>	lə <sup>3</sup>	yə <sup>1</sup>	yə <sup>1</sup>	lə <sup>3</sup>
213	today	zare	harra	ək <sup>w</sup> a <sup>1</sup>	alkka <sup>1</sup>	əkk <sup>w</sup> a <sup>1</sup>	ək <sup>w</sup> a <sup>1</sup>	ək <sup>w</sup> a <sup>1</sup>	ok <sup>w</sup> a <sup>1</sup>	axum <sup>1</sup>	ak <sup>w</sup> a <sup>1</sup>	əkk <sup>w</sup> a <sup>1</sup>	awgə <sup>2</sup>
214	tomato	timatim <sup>1</sup>	timaatim <sup>1</sup>	timatim <sup>1</sup>	timatim <sup>1</sup>	timatim <sup>1</sup>	timatim <sup>1</sup>	timatim <sup>1</sup>	timatim <sup>1</sup>	tuntum <sup>1</sup>	timatim <sup>1</sup>	timatim <sup>1</sup>	timatim <sup>1</sup>
215	tomorrow	nəgə <sup>1</sup>	boru	nəgə <sup>1</sup>	nəgə <sup>1</sup>	nəgə <sup>1</sup>	nəgə <sup>1</sup>	nəgə <sup>1</sup>	nəgə <sup>1</sup>	nəgə <sup>1</sup>	nəgə <sup>1</sup>	nəgə <sup>1</sup>	gees <sup>2</sup>
216	too	bəfam <sup>3</sup>	baay'ee	nəkkar <sup>1</sup>	haməʔad <sup>2</sup>	nəkkar <sup>1</sup>	nəkkar <sup>1</sup>	nəkkar <sup>1</sup>	buseʔəha <sup>2</sup>	burim <sup>3</sup>	bəfam <sup>3</sup>	likkə4	sirəm <sup>5</sup>
217	traveler	məngədəni <sup>3</sup>	eemala <sup>1</sup>	eməna <sup>1</sup>	meñña <sup>1</sup>	eməna <sup>1</sup>	eməna <sup>1</sup>	eməna <sup>1</sup>	eməna <sup>1</sup>	əngədəni <sup>3</sup>	imatəni <sup>3</sup>	məngədəni <sup>3</sup>	ungawəro <sup>4</sup>
218	trivet	gulləcca <sup>3</sup>	sunsuma	gonziyə <sup>1</sup>	gamziyə <sup>1</sup>	gamziyə <sup>1</sup>	gulləcca <sup>3</sup>	g <sup>w</sup> amziyye <sup>1</sup>	midacca <sup>2</sup>	gawziyə <sup>1</sup>	akufa <sup>4</sup>	gonziyə <sup>1</sup>	šitkar <sup>5</sup>
219	trouser	sutri <sup>2</sup>	kofoo	bolale <sup>1</sup>	sureʔ <sup>2</sup>	sulle <sup>2</sup>	bolale <sup>1</sup>	bolale <sup>1</sup>	bolale <sup>1</sup>	bolale <sup>1</sup>	luxe <sup>2</sup>	bolale <sup>1</sup>	bolale <sup>1</sup>
220	two	hulət <sup>2</sup>	lama	x <sup>w</sup> et <sup>1</sup>	hurʔət <sup>1</sup>	x <sup>w</sup> ett <sup>1</sup>	x <sup>w</sup> et <sup>1</sup>	x <sup>w</sup> et <sup>1</sup>	wirʔet <sup>1</sup>	kitt <sup>1</sup>	x <sup>w</sup> ett <sup>1</sup>	x <sup>w</sup> ett <sup>1</sup>	ošt <sup>2</sup>
221	upon	ilayulay	gubbaarra	bəforəta <sup>1</sup>	bəforfor <sup>1</sup>	bəforəta <sup>1</sup>	bəforəta <sup>1</sup>	bəforəta <sup>1</sup>	bəforuha <sup>1</sup>	bəlaləh <sup>w</sup> an <sup>2</sup>	bəforəhuf <sup>w</sup> ər <sup>1</sup>	bəf <sup>w</sup> eluta <sup>1</sup>	bəddərka <sup>3</sup>
222	us	imnan <sup>1</sup>	nu	yina <sup>1</sup>	inə <sup>2</sup>	yina <sup>1</sup>	yina <sup>1</sup>	ina <sup>1</sup>	ina <sup>1</sup>	ina <sup>1</sup>	imna <sup>1</sup>	imnan <sup>1</sup>	yəñə <sup>4</sup>
223	waist-band	məkənət	sabbata	azgart <sup>1</sup>	azgatta <sup>1</sup>	azgard <sup>1</sup>	azgart <sup>1</sup>	azgart <sup>1</sup>	azgat <sup>1</sup>	dəkot <sup>2</sup>	azlagidd <sup>1</sup>	azgalidd <sup>1</sup>	imro <sup>3</sup>
224	war	fər	waraana	gaz <sup>1</sup>	ariv <sup>2</sup>	arb <sup>2</sup>	gaz <sup>1</sup>	gaz <sup>1</sup>	gaz <sup>1</sup>	tiğə <sup>3</sup>	arib <sup>2</sup>	gaz <sup>1</sup>	arb <sup>2</sup>
225	was	nəbbər <sup>2</sup>	ture	banə <sup>1</sup>	banədə <sup>1</sup>	banəmə <sup>3</sup>	banə <sup>1</sup>	banə <sup>1</sup>	banədə <sup>3</sup>	nəbbərə <sup>3</sup>	banə <sup>1</sup>	banno <sup>1</sup>	narə <sup>2</sup>
226	water	wiha	bišaan	ixə <sup>1</sup>	ixə <sup>1</sup>	ixə <sup>1</sup>	ixə <sup>1</sup>	ixə <sup>1</sup>	ixə <sup>1</sup>	yiga <sup>1</sup>	əga <sup>1</sup>	iga <sup>1</sup>	məy <sup>2</sup>
227	wednesday	irob <sup>1</sup>	roobii <sup>1</sup>	əro <sup>1</sup>	arbe <sup>1</sup>	ərow <sup>1</sup>	əro <sup>1</sup>	əro <sup>1</sup>	xarob <sup>1</sup>	əro <sup>1</sup>	əro <sup>1</sup>	əro <sup>1</sup>	harbe <sup>1</sup>
228	were	nəbbər <sup>2</sup>	turan	banəwo <sup>1</sup>	banowətə <sup>1</sup>	banəbo <sup>1</sup>	banəbo <sup>1</sup>	banəwo <sup>1</sup>	banowətə <sup>1</sup>	nəbbərim <sup>2</sup>	banno <sup>1</sup>	banimutt <sup>1</sup>	naru <sup>2</sup>
229	what	min <sup>1</sup>	maal <sup>1</sup>	mirkar <sup>1</sup>	mirafor <sup>1</sup>	məʔe <sup>1</sup>	mirkar <sup>1</sup>	mirkar <sup>1</sup>	miraʔar <sup>1</sup>	min <sup>1</sup>	mirkar <sup>1</sup>	miraʔe <sup>1</sup>	min <sup>1</sup>
230	wheat	sinde <sup>1</sup>	kamadi	sina <sup>1</sup>	sineʔ <sup>1</sup>	sinne <sup>1</sup>	sinay <sup>1</sup>	sine <sup>1</sup>	sinaʔ <sup>1</sup>	sirre <sup>1</sup>	sirre <sup>1</sup>	sirre <sup>1</sup>	siraʔ <sup>1</sup>
231	which	yətinihaw	kam	əndehuta <sup>2</sup>	ətədə <sup>1</sup>	indeyətə <sup>1</sup>	etətə <sup>1</sup>	etətə <sup>1</sup>	ətədə <sup>1</sup>	yitta <sup>1</sup>	eti <sup>1</sup>	im <sup>w</sup> ewe <sup>1</sup>	aytay <sup>1</sup>
232	who	man <sup>1</sup>	eeñnu	m <sup>w</sup> an <sup>1</sup>	man <sup>1</sup>	m <sup>w</sup> an <sup>1</sup>	m <sup>w</sup> an <sup>1</sup>	m <sup>w</sup> an <sup>1</sup>	m <sup>w</sup> an <sup>1</sup>	man <sup>1</sup>	ma <sup>1</sup>	m <sup>w</sup> a <sup>1</sup>	maani <sup>1</sup>
233	whoever	maniñhawm <sup>2</sup>	kamiyyu	m <sup>w</sup> animənda <sup>2</sup>	attimanir <sup>1</sup>	indeyənda <sup>2</sup>	etətə <sup>1</sup>	etənda <sup>1</sup>	itindra <sup>1</sup>	yittañña <sup>1</sup>	etəna <sup>1</sup>	mannimina <sup>2</sup>	aynaannə <sup>3</sup>
234	wide (adj)	səffi	balʔa	bətit <sup>1</sup>	bəttəd <sup>1</sup>	bəttit <sup>1</sup>	bətit <sup>1</sup>	bətit <sup>1</sup>	bətit <sup>1</sup>	bətat <sup>1</sup>	bəttat <sup>1</sup>	bəttit <sup>1</sup>	fətt <sup>2</sup>
235	wind	nifas <sup>1</sup>	killeensa	infas <sup>1</sup>	infas <sup>1</sup>	nəfas <sup>1</sup>	nifas <sup>1</sup>	infas <sup>1</sup>	nəfas <sup>1</sup>	nəfas <sup>1</sup>	nifas <sup>1</sup>	infas <sup>1</sup>	nifas <sup>1</sup>
236	with	gar <sup>1</sup>	wəggūm	tahut <sup>2</sup>	gi <sup>1</sup>	tihuda <sup>2</sup>	gam <sup>w</sup> a <sup>1</sup>	gam <sup>w</sup> a <sup>1</sup>	gəz <sup>w</sup> a <sup>1</sup>	goy <sup>1</sup>	təge <sup>1</sup>	ge <sup>1</sup>	tuuha <sup>2</sup>
237	woman	setiyyo	gəarti	mišt <sup>1</sup>	mist <sup>1</sup>	mišt <sup>1</sup>	mišt <sup>1</sup>	mišt <sup>1</sup>	mišt <sup>1</sup>	mišt <sup>1</sup>	mišt <sup>1</sup>	mišt <sup>1</sup>	mišt <sup>1</sup>
238	word	kal <sup>1</sup>	gəcča	kār <sup>1</sup>	kāl <sup>1</sup>	kār <sup>1</sup>	kār <sup>1</sup>	kār <sup>1</sup>	kār <sup>1</sup>	kāl <sup>1</sup>	kāl <sup>1</sup>	kāl <sup>1</sup>	kāl <sup>1</sup>
239	wrapped (adj)	fəkkil <sup>1</sup>	maramaa	fəkkil <sup>1</sup>	dububil <sup>2</sup>	fəkkil <sup>1</sup>	fəbarir <sup>1</sup>	dibibil <sup>2</sup>	fəkkil <sup>1</sup>	dibibil <sup>2</sup>	fəkkil <sup>1</sup>	fəkkil <sup>1</sup>	fəkkil <sup>1</sup>
240	you (2.M.SG)	antə <sup>1</sup>	ati <sup>1</sup>	axə <sup>1</sup>	axa <sup>1</sup>	zahə <sup>1</sup>	axə <sup>1</sup>	axə <sup>1</sup>	axə <sup>1</sup>	dəxa <sup>1</sup>	axa <sup>1</sup>	atə <sup>2</sup>	atə <sup>2</sup>

## B.2 List of words for Word Categorization test

The following list of words were used in the Semantic Word categorization test to measure the degree of intelligibility and to determine the functional distance among the selected language varieties.

<b>cloths</b>	<b>Body parts</b>	<b>Kitchen utilities</b>	<b>Fruits</b>	<b>Food type</b>
shoes	finger	spoon	banana	bread
shirt	lip	ladle	mango	'kocho'
hat	eye	pan	orange	'injera'
belt	arm	knife	berry	stew
trouser	breast	cutting board	guava	pancake
handkerchief	leg	griddle	cherimoya	roasted meat
dress	chest	stirring rod	coke	mush
shorts	eye	kettle	tangerine	'besso'
waist-band	hair	food-table	lemon	porridge
headdress	neck	plate	doviyalis abyssnica	roasted grain
<b>Domestic animal</b>	<b>Furniture</b>	<b>Vegetables</b>	<b>Wild animal</b>	<b>Cereals</b>
hen	table	cabbage	elephant	barely
ox	chair	pepper	lion	wheat
camel	shelf	tomato	tiger	maize
donkey	locker	onion	hyena	pea
goat	bed	potato	crocodile	fava bean
sheep	sofa	carrot	giraffe	sorghum
dog	stool	garlic	monkey	'teff'
cat	chassis	pumpkin	ape	bean
horse	mirror	sweet potato	fox	lentil
mule	box	basil	gazelle	chickpea



### B.3 Word Categorization, word order

**Description:** words used for word categorization were listed in different orders to block the priming effect. Different CDs were created by changing the order of the varieties in which the words are spoken. One CD was used for one language area. Each CD consists of ten tracks. One participant matched 10 list of words within a track with their semantic categories provided on the answer sheet. Matching the full CD requires the involvement 10 participants. In our case, each CD was repeated 3 times and administered to the total of around 30 students.

	CDs	CD1	CD2	CD3	CD4	CD5	CD6	CD7	CD8	CD9	CD10
No	Words	TR1	TR1	TR1	TR1	TR1	TR1	TR1	TR1	TR1	TR1
1	shoe	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan	Gura	Gumer	Ezha	Inor
2	finger	Inor	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan	Gura	Gumer	Ezha
3	spoon	Ezha	Inor	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan	Gura	Gumer
4	banana	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan	Gura
5	bread	Gura	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan
6	hen	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane	Silte	Muher
7	table	Muher	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane	Silte
8	cabbage	Silte	Muher	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane
9	elephant	Kistane	Silte	Muher	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn	Chaha
10	barley	Chaha	Kistane	Silte	Muher	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn
		TR2	TR2	TR2	TR2	TR2	TR2	TR2	TR2	TR2	TR2
11	lip	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan	Gura	Gumer	Ezha	Inor
12	ladle	Inor	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan	Gura	Gumer	Ezha
13	mango	Ezha	Inor	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan	Gura	Gumer
14	koc'c'o	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan	Gura
15	Ox	Gura	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan
16	chair	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane	Silte	Muher
17	pepper	Muher	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane	Silte
18	lion	Silte	Muher	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane
19	wheat	Kistane	Silte	Muher	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn	Chaha
20	shirt	Chaha	Kistane	Silte	Muher	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn
		TR3	TR3	TR3	TR3	TR3	TR3	TR3	TR3	TR3	TR3
21	pan	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan	Gura	Gumer	Ezha	Inor
22	orange	Inor	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan	Gura	Gumer	Ezha
23	inǧora	Ezha	Inor	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan	Gura	Gumer
24	camel	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan	Gura
25	shelf	Gura	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan
26	tomato	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane	Silte	Muher
27	tiger	Muher	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane	Silte
28	maize	Silte	Muher	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane
29	hat	Kistane	Silte	Muher	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn	Chaha
30	eye	Chaha	Kistane	Silte	Muher	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn
		TR4	TR4	TR4	TR4	TR4	TR4	TR4	TR4	TR4	TR4
31	berry	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan	Gura	Gumer	Ezha	Inor
32	type of stew	Inor	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan	Gura	Gumer	Ezha



33	donkey	Ezha	Inor	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan	Gura	Gumer
34	locker	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan	Gura
35	onion	Gura	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan
36	hyena	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane	Silte	Muher
37	pea	Muher	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane	Silte
38	belt	Silte	Muher	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane
39	am	Kistane	Silte	Muher	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn	Chaha
40	knife	Chaha	Kistane	Silte	Muher	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn
		TR5	TR5	TR5	TR5	TR5	TR5	TR5	TR5	TR5	TR5
41	pancake	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan	Gura	Gumer	Ezha	Inor
42	goat	Inor	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan	Gura	Gumer	Ezha
43	bed	Ezha	Inor	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan	Gura	Gumer
44	potato	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan	Gura
45	crocodile	Gura	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan
46	fava bean	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane	Silte	Muher
47	trouser	Muher	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane	Silte
48	breast	Silte	Muher	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane
49	cutting board	Kistane	Silte	Muher	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn	Chaha
50	guava	Chaha	Kistane	Silte	Muher	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn
		TR6	TR6	TR6	TR6	TR6	TR6	TR6	TR6	TR6	TR6
51	sheep	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan	Gura	Gumer	Ezha	Inor
52	sofa	Inor	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan	Gura	Gumer	Ezha
53	carrot	Ezha	Inor	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan	Gura	Gumer
54	giraffe	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan	Gura
55	sorghum	Gura	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan
56	handkerchief	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane	Silte	Muher
57	leg	Muher	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane	Silte
58	griddle	Silte	Muher	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane
59	chenmoya	Kistane	Silte	Muher	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn	Chaha
60	roasted meat	Ch's aha	Kistane	Silte	Muher	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn
		TR7	TR7	TR7	TR7	TR7	TR7	TR7	TR7	TR7	TR7
61	stool	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan	Gura	Gumer	Ezha	Inor
62	garlic	Inor	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan	Gura	Gumer	Ezha
63	monkey	Ezha	Inor	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan	Gura	Gumer
64	'teff'	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan	Gura
65	females' dress	Gura	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan
66	chest	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane	Silte	Muher
67	stirring rod	Muher	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane	Silte
68	coke	Silte	Muher	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane
69	mush	Kistane	Silte	Muher	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn	Chaha
70	dog	Chaha	Kistane	Silte	Muher	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn

		TR8	TR8	TR8	TR8	TR8	TR8	TR8	TR8	TR8	TR8
71	pumpkin	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan	Gura	Gumer	Ezha	Inor
72	ape	Inor	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan	Gura	Gumer	Ezha
73	bean	Ezha	Inor	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan	Gura	Gumer
74	shorts	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan	Gura
75	Eye	Gura	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan
76	kettle	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane	Silte	Muher
77	tangerine	Muher	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane	Silte
78	'bḗḗḗḗ'	Silte	Muher	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane
79	cat	Kistane	Silte	Muher	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn	Chaha
80	chassis	Chaha	Kistane	Silte	Muher	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn
		TR9	TR9	TR9	TR9	TR9	TR9	TR9	TR9	TR9	TR9
81	fox	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan	Gura	Gumer	Ezha	Inor
82	lentils	Inor	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan	Gura	Gumer	Ezha
83	waist-band	Ezha	Inor	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan	Gura	Gumer
84	hair	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan	Gura
85	food-table	Gura	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan
86	lemon	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane	Silte	Muher
87	porridge	Muher	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane	Silte
88	horse	Silte	Muher	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane
89	mirror	Kistane	Silte	Muher	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn	Chaha
90	sweet potato	Chaha	Kistane	Silte	Muher	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn
		TR10	TR10	TR10	TR10	TR10	TR10	TR10	TR10	TR10	TR10
91	chick-pea	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan	Gura	Gumer	Ezha	Inor
92	headdress	Inor	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan	Gura	Gumer	Ezha
93	neck	Ezha	Inor	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan	Gura	Gumer
94	Plate	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan	Gura
95	doviyahis abysimica	Gura	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane	Silte	Muher	Mesqan
96	roasted grain	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn	Cheha	Kistane	Silte	Muher
97	mule	Muher	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane	Silte
98	box	Silte	Muher	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn	Chaha	Kistane
99	basil	Kistane	Silte	Muher	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn	Chaha
100	gazelle	Chaha	Kistane	Silte	Muher	Mesqan	Gura	Gumer	Ezha	Inor	Endegagn

---

# Appendix C

## Supplementary Results

### C.1 Phonetic similarity index

Language	CH	EN	EZ	GM	GU	IN	KS	MS	MU	SI
Chaha	100									
Endegagn	82	100								
Ezha	92	81	0							
Gumer	92	82	92	100						
Gura	95	82	90	93	100					
Inor	88	86	85	86	88	100				
Kistane	82	79	82	81	81	79	100			
Mesqan	89	80	88	87	88	83	87	100		
Muher	86	79	87	85	85	81	88	88	100	
Silt'e	80	77	79	78	80	76	82	81	78	100

### C.2 Lexical similarity index

Language	CH	ED	EZ	GM	GU	IN	KS	MS	MU	SI
Chaha	0									
Endegagn	74	100								
Ezha	87	75	0							
Gumer	88	73	87	100						
Gura	89	73	85	88	100					
Inor	78	82	77	80	80	100				
Kistane	61	59	66	67	63	63	100			
Mesqan	76	69	78	79	76	72	70	100		
Muher	79	67	80	80	76	72	72	82	100	
Silt'e	52	50	53	52	51	53	56	54	53	100

### C.3 Perceptual similarity index

Language	CH	EN	EZ	GM	GU	IN	KS	MS	MU	SI
Chaha	81	13	52	95	96	38	12	23	29	2
Endegagn	37	100	44	49	51	91	57	51	42	34
Ezha	86	10	89	93	90	27	25	66	51	7
Gumer	78	19	70	95	92	40	20	54	43	10
Gura	84	35	72	85	89	57	18	47	37	10
Inor	47	72	36	60	55	99	25	44	36	8
Kistane	30	33	30	21	23	33	99	54	76	20
Mesqan	38	21	63	59	71	17	58	96	78	11
Muher	58	10	35	72	72	22	40	59	96	13
Silt'e	23	32	35	33	30	20	63	52	45	100

### C.4 Attitude test results

Language	ED	IN	EZ	GM	GU	MS	MU	SI	KS	CH
Endegagn	99	93	47	61	60	53	61	42	58	48
Inor	75	99	49	76	72	48	41	23	24	54
Ezha	19	25	84	91	88	55	50	11	29	81
Gumer	40	44	64	92	87	40	42	17	21	74
Gura	31	48	65	79	87	33	21	07	09	75
Mesqan	34	33	62	58	65	91	73	36	66	48
Muher	28	41	48	82	81	49	95	33	48	74
Silt'e	50	39	45	49	45	53	52	95	68	34
Kistane	49	52	48	46	43	65	73	38	91	46
Chaha	10	32	46	96	96	18	25	03	10	81

## C.5 Cophenetic distance among the nodes

No	Pairs of Varieties	Combined	Hetzron (1972)	Demeke (2001)
1	CH-EN	5	4	4
2	CH-EZ	2	2	2
3	CH-GM	2	2	2
4	CH-GU	2	2	2
5	CH-IN	5	4	2
6	CH-KS	5	8	7
7	CH-MS	4	4	6
8	CH-MU	4	5	2
9	CH-SI	6	10	8
10	EZ-EN	4	4	3
11	EZ-IN	5	4	2
12	GM-EN	4	4	3
13	GM-EZ	2	2	2
14	GM-IN	5	4	2
15	GU-EN	5	4	3
16	GU-EZ	2	2	2
17	GU-GM	2	2	2
18	GU-IN	5	4	2
19	IN-EN	2	2	3
20	KS-EN	5	8	7
21	KS-EZ	5	8	7
22	KS-GM	5	8	7
23	KS-GU	4	8	7
24	KS-IN	5	8	7
25	KS-MS	5	6	3
26	KS-MU	5	7	7
27	KS-SI	3	8	9
28	MS-EN	5	4	6
29	MS-EZ	4	4	6
30	MS-GM	4	5	6
31	MS-GU	4	4	6
32	MS-IN	5	4	6
33	MU-EN	5	5	4
34	MU-EZ	4	5	2
35	MU-GM	4	5	2
36	MU-GU	4	4	2
37	MU-IN	5	5	2
38	MU-MS	2	3	4
39	SI-EZ	6	10	8
40	SI-GM	6	10	8
41	SI-GU	6	10	8
42	SI-IN	6	10	8
43	Silté-EN	6	10	8
44	SI-MS	5	8	8
45	SI-MU	6	8	8

Figure C.1: Cophenetic distance



## C.6 Results of fuzzy clustering

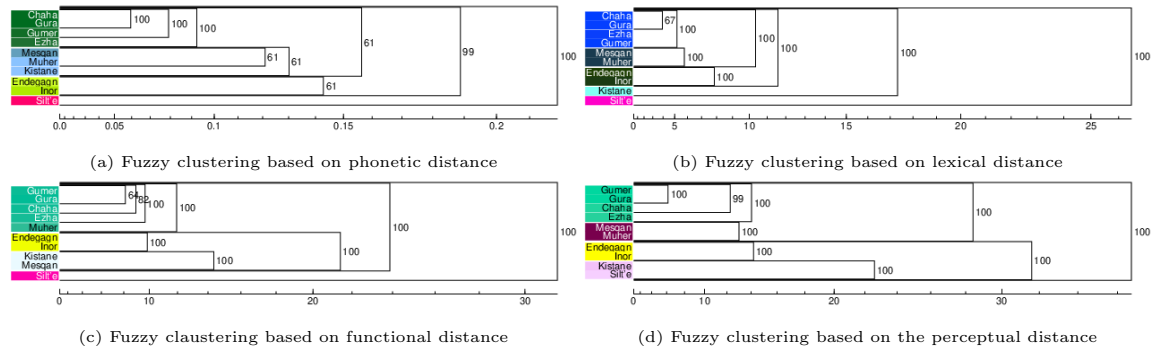


Figure C.2: Fuzzy clustering based on the structural, functional and perceptual distances

## C.7 Test participants

Note that the number of participants of the perception test and the mutual intelligibility is slightly different. This is because, the two tests were administered at different times in some of the areas, and the respondents who were absent on one of the tests can be present on the other.

### C.1. Participants of the word categorization test

The table contains CD number administered in each area (CD No.), the range of age of the participants (Age), the participants' grade level (G.9 = Grade nine, G.10 = Grade 10, G. 11 = Grade 11, G. 12 = Grade 12), the participants whose responses were considered for the analysis (considered), the participants whose responses were not considered (excluded- because sometimes some respondents stacked on the first page of the response sheet while the CD was playing).

Area	CD No.	Age	Sex		Grade				Total	Considered	Excluded
			M	F	G.9	G.10	G.11	G.12			
Gura	01	15-22	23	6	1	6	7	15	29	29	-
Gumer	02	17-21	17	12	-	-	29	-	29	28	1
Inor	03	16-21	24	5	-	29	-	-	29	22	7
Chaha	04	16-23	18	11	-	-	15	11	29	26	3
Muher	05	18-21	14	15	-	-	17	12	29	24	5
Ezha	06	16-23	14	16	-	-	9	21	30	26	4
Endogagn	07	18-24	21	7	-	-	22	6	28	22	6
Kistane	08	17-22	20	10	-	-	30	-	30	23	7
Mesqan	09	16-21	8	21	12	19	-	-	29	25	4
Silt'e	10	18-24	12	11	-	-	-	23	23	23	-
			171	114	13	54	129	88	285	248	37

### C.2. Participants of the Perceptual Test

The participants who did not follow the instruction properly were excluded.

Area	Age	Sex		Grade				Total	Considered	Excluded
		M	F	G.9	G.10	G.11	G.12			
Gura	15-22	23	5	1	7	7	15	30	28	2
Gumer	17-21	15	14	-	-	30	-	30	25	5
Inor	16-22	21	7	-	28	-	-	28	20	8
Chaha	16-23	18	10	-	-	17	11	28	25	3
Muher	18-21	15	15	-	-	18	12	30	26	4
Ezha	16-23	16	14	-	-	10	20	30	27	3
Endogagn	18-24	21	7	-	-	23	5	28	28	-
Kistane	17-23	19	11	-	-	30	-	30	30	-
Mesqan	16-21	8	21	12	19	-	-	29	25	4
Silt'e	18-23	15	14	-	-	-	29	29	24	5
Total		171	118	13	54	135	92	289	258	34